

Muddle over superhighway builders

The British government is heading for a monumental muddle in attempting to decide which monopoly should take charge of any British information superhighway there may be. Other governments should watch what happens.

WILL the British government or its successor (expected in 1996) be compelled to renationalize British Telecom, the largest communications company in Britain and the privatized successor to what was once a public monopoly? The question is not political, but technical and administrative. The British government (whichever party forms it) is heading for an impossible muddle over telecommunications. Faced with the difficulty of making sense of the patchwork it has made for itself, it may well decide that it has no choice but to start again, with a blank sheet of paper. But nobody can be sure that it would make a better job of shaping the telecommunications industry if it had a second chance.

The circumstances are these. The privatization of British Telecom was at the centre of the ambitions of the Conservative government elected 15 years ago. Soon afterwards, in the person of Mr Kenneth Baker, then minister of information technology, the government also decided to "cable Britain", dividing Britain into geographical pieces and awarding franchises to companies willing to provide cable-television services commercially. The franchises required that the operators should install cable systems that were technically advanced in 1981. The cabling of Britain has nevertheless hung fire, partly because of the tax treatment of the money spent in digging up Britain's streets. To make the packages more palatable, the cable companies have been allowed to offer telephony as well as television signals to their subscribers, but British Telecom cannot (for another two years) offer its subscribers even a pay-TV service.

Now, matters have been further complicated by the advocacy by the US vice-president, Al Gore, of the 'information superhighway'. Who would build its British equivalent? A House of Commons select committee is expected this week to report that the cost would amount to £15 billion and that the benefits would be worth much more. That is believable. But the essence of the multi-media business is that offices and houses are connected by a single broad-bank link to a network carrying a variety of signals.

The dilemma for the British government is that it will soon be forced to choose whose holes in the ground should carry the optical-fibre cables of the superhighway, British Telecom's or the cable companies? Either choice would put the neglected competitor out of business, causing a huge rumpus. The shareholders would have to be compensated, perhaps even bought out. Requiring British Telecom to buy out the cable companies would be another solution, but one that would saddle consumers with needless costs and recre-

ate a monopoly in telecommunications the present government hoped to end for good.

The moral in this tale is simple. Arranging that monopolies, whether public or private, function efficiently and deal equitably with their customers is never easy. Doing so in fields in which the technology is changing quickly is especially dangerous, and can be disastrous. In the British case, short-sighted franchising has already saddled the country with almost twice as much cabling beneath the streets as it really needs, together with an organizational structure that cannot easily be used to operate an information superhighway. Starting all over again might well be simpler. □

Goodbye to academia

US grant agencies should be kinder to postdocs or reckon that the supply of young researchers will shrink.

Is research no longer attractive to the young? That is the question provoked by the surprising decline in the number of research grants to people under 37 awarded by the US National Institutes of Health (NIH). By all accounts, the absolute number of grants to people in this age-cohort fell by 54 per cent between 1985 and 1993 (see *Nature* 370, 87; 1994). Even if the cornucopia that is NIH may not be what it was when its budget grew much faster than inflation, it is natural that the managers are asking, "Where have all the young researchers gone?"

The managers will find parts of an explanation under their own noses, in the competitive grants system they administer. Over three decades, this system has endowed the United States with the most competitive and productive system of biomedical research the world has ever known. Principal investigators compete for grants in the knowledge that only success can ensure that they and their colleagues continue to work in an academic environment. It is a commonplace that the burden of competition falls on the shoulders of the army of postdoctoral fellows inhabiting US biomedical research laboratories; theirs are the lights routinely burning until the small hours of the morning, accumulating the data on which the next grant application will be based, or slicing a publishable paper into smaller pieces to magnify the laboratory's publications record. They are also the most vulnerable members of the team, the first to lose their jobs when a grant comes to an end or is curtailed. Rarely will a parent institution help



them to bridge even a short spell on the streets.

It is a marvellously productive yet cruel system, comparable with child labour in both respects. The difference is that the victims of the postdoctoral system in the United States are no longer children, but grown men and women, often with families and family responsibilities. Now that the traditional reward for consecutive spells as a postdoctoral fellow — a tenured position at some academic institution — is itself receding, it would be only natural if trained researchers let their ambitions embrace other goals, in industry or even commerce. The remedy is one in any case required by natural justice: a more civilized way of dealing with the careers of the non-commissioned officers of research. Thirty-seven, after all, is hardly a reasonable definition of youth. Einstein had done all his important work by that age. □

Regulating embryology

A premature ban on the use of fetal oocytes in the treatment of infertility is a needless mistake.

SHOULD human oocytes taken from an aborted human fetus be used, after fertilization, in the treatment of infertility? The answer is unambiguous, and "NO!". Despite the present shortage of eggs for use in infertility treatment, the use of eggs from aborted fetuses would be improper while so little is known of the reasons why the number of surviving oocytes in the human ovary declines so drastically (by a factor of 100,000) between birth and adulthood. Is this a random process designed to reduce the metabolic cost of being female, or nature's built-in mechanism for genetic screening and the elimination of disadvantageous eggs? While that question is undecided, the use of fetal oocytes would be potentially dangerous, and therefore wrong. And some time must pass before there is an answer.

That is part of the background to the publication last week of a new policy statement by Britain's hyperactive Human Fertilization and Embryology Authority, which has statutory responsibility for licensing fertility treatment (including artificial insemination) and research. The new statement deals with the possible use of eggs from three sources: aborted female fetuses, the cadavers of adult women and the ovaries of women who are alive and who volunteer to donate ovarian tissue (see page 241). On the last, the committee asks only that there should be full informed consent (which entails that the donors should be adult), on the second, it advocates a system of donor cards (as for other organs) which has not yet been developed, but it then goes on to declare a ban on the use of fetal oocytes except in research.

These policy decisions are sheer busybodyness. The use of adult ovaries in the manner the committee recommends is covered by existing legislation (including that under which the committee itself exists). The committee did not need a "policy decision" to regulate them. It would also have been well within its present competence to have refused licences to those seeking to use fetal oocytes for the treatment of infertility until there is a better understanding of the biology

of the natural attenuation of the oocyte population. That, indeed, would have been the time to have launched the consultation exercise on which the committee has laboured for six months, and whose most memorable product appears to have been a write-in campaign by anti-abortionists. The danger in this needless fuss is that the committee's claim on seriousness will be undermined (see *Nature* 361, 283; 1993). □

Eastern jobs go west

German reunification continues to exact a high price in the east, but one that may be worth paying.

GERMANY is paying a high price for unification, not only through the government's increased load of public debt, but in the repeated reminders to those in the six new *Länder* of the east that the old German Democratic Republic collapsed into the arms of West Germany three years ago on terms determined largely in the west. The reorganization of eastern universities on western lines was one of the conditions of reunification. There have been just complaints at the tests of political correctness to which eastern academics were subjected. Now, many east German academics have lost their jobs in open competition with people they will regard as carpetbaggers from the west (see page 240).

The hope must be that this latest humiliation will not be the divisive influence it may seem. Although the eastern economy still lags behind that of western Germany, Germany is now a single political entity, with all that implies for the present and the future. A generation from now, there is every chance that people will have forgotten the old division.

In any case, the competition for university posts appears to have been conducted on merit only. Those who have lost their posts will not lose their pension rights. And in the long run it will help the east that a number of academics from western Germany have thrown in their lot with eastern universities. (They are far outnumbered by the commercially minded people seeking to make their fortunes to the east.) Whatever their strengths in scholarship and research, universities in the east can only benefit from a modest influx of people used to the argumentative and free-thinking ways of the traditional German university. And while the numbers may be small, go-getting academics from the west could help to unlock sources of research funds not yet tapped. □

One-day conferences

A *Nature* meeting, *Genetics in Europe* now, will be held in Berlin on Friday 30 September 1994 at the Brandenburg Akademie, and the cost is DM195 or £75 (plus VAT).

A meeting in Paris, *What is distinctive about European science?*, will be held on Wednesday 26 October 1994 at the Ecole Normale Supérieure. The cost is FF675 or £75 (plus VAT).

Payment may be made by cheque (payable to Macmillan Subscriptions Ltd) or credit card. For further details of either meeting, or to reserve places, please write, telephone or fax to Pippa Burnett, Nature Conferences Office, 4 Little Essex Street, London WC2R 3LF, UK; telephone +44 71 836 6633 ext. 2593; fax +44 71 379 5417.

UK urged to emulate US in linking military research to civilian goals

London. Britain's Ministry of Defence (MoD) has been urged to allow its research laboratories to play a far greater role than they can at present in contributing to the nation's economic prosperity, and not merely to its physical defence.

The proposal comes from a committee of the House of Lords, which quotes with approval the way the United States has included that concept of economic security within the broader goals of national security. The committee points, for example, to the Defense Advanced Research Projects Agency (DARPA) — as well as to comparable developments in France — as an example that the United Kingdom could profitably follow.

But the committee also acknowledges that achieving this would require changes in the administration of defence research: these range from revisions in the Treasury rules under which defence laboratories operate, through persuading the Department of Trade and Industry (DTI) to boost funding for joint research projects, to giving the Office of Science and Technology (OST) a more direct say in the planning of the laboratories' research programmes.

Given that each would meet substantial political opposition — the MoD for example, has already fought keenly (and successfully) against the transfer of any control over its resources to the OST — the government

is unlikely to accept the broad thrust of the proposals with open arms.

The recommendations are made by the House of Lords's Select Committee on Science and Technology. They are based on an inquiry into the functioning of the Defence Research Agency (DRA), whose report was published this week.

Until recently, MoD laboratories operated under budgets allocated centrally from the ministry. Since 1991, however, those responsible for all research except for that on nuclear, chemical and biological weapons, have operated under the umbrella of the DRA, and have been required to raise their funds through separate contracts with individual 'customers' within the defence establishment.

"In general, we felt that the DRA has been doing a good job," says the committee's chairman, Lord Selborne, pointing in particular to the increased cost-effectiveness of its individual establishments, and the relevance of their work to specific defence needs. But, he adds, "we do not feel that it [the DRA] has had a fair deal from the MoD itself", either in being allowed to plan for its own long-term future — or in maximizing its potential contribution to the civilian economy.

One problem is the difficulty faced by the DRA as a 'contractor' organization in developing a long-term view of future re-

search needs. Indeed, the committee points out that even the DRA, in a corporate plan that is not publicly available, says it is an "illusion" to believe that its business planning can be based on the MoD's long-term costing process.

Selborne points out that several witnesses to the inquiry expressed fears that the DRA will be forced by the need for greater commercial discipline to focus on more short-term research and development at the expense of its longer-term research. "We accept that these fears may be well-founded," the committee's report says.

Furthermore, requiring the DRA to seek a customer for every activity "does not create an environment in which DRA can readily plan ahead and establish a strong foothold in support of wealth creation beyond the defence sector".

Finally, the committee suggests that the OST should be given far greater power to ensure that the MoD plays an active role in allowing the DRA's research facilities to be used for the broader goal of wealth creation — and not confined, as at present, to meeting the technological needs of the armed forces.

One solution, it suggests, could be to give OST (as well as the DTI) part ownership of DRA. Another would be to give the government's chief scientific adviser the authority "to expose any links which impeded closer links between civil and military science and technology".

Defence administrators, like other civil servants, are fiercely protective of their turf, and are unlikely to view either proposal with much enthusiasm. But Selborne points out that, whatever reservations they may have, the proposals are in line with the main thrust of last year's white paper on science, *Realizing our Potential*.

Indeed, Selborne implies that, unless the OST is able to flex its muscles in this or similar ways, it will have failed to meet expectations raised by the white paper that the government is genuinely committed to focusing the nation's whole research base on the twin needs of wealth creation and improving the quality of life.

Under the new arrangements for defence research announced last week, the DRA will become part of a new Defence Science Technology Agency. The government also says that it intends to separate 'corporate' and 'contract' research, with the implication that the first of these will include responsibility for meeting long-term scientific and technical needs. But it has also announced that a key goal will be to reduce spending on research in both categories.

David Dickson

New faces to head British science

London. Last week's reshuffle of the British cabinet saw the arrival of a new science minister with the appointment of David Hunt, formerly employment minister, to the position of Chancellor of the

Duchy of Lancaster, responsible for the Office of Public Services and Science.

The permanent secretary in the office is Robert Hughes, formerly a parliamentary whip for the Conservative Party (and, before his election to Parliament in 1987, a news editor with the British Broadcasting Corporation).

The fact that Hunt has been given wider-ranging political responsibilities than his predecessor, William Waldegrave — he will, for example,

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Robert Hughes (left) and David Hunt

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

himself has been quick to counter any impression that science has been downgraded, saying that his experience at the Department of Employment "has shown me how important science is to wealth creation".

At the same time, however, Cabinet officials say that day-to-day responsibility for the activities of the Office of Science and Technology will be taken on by Hughes, who will act as the "junior science minister". □

oversee the activities of various Cabinet committees — has led some observers to express concern that the reshuffle may result in science slipping down the political priority list.

But Hunt

'Reasonable pricing' may be dropped from NIH deals

Washington. The US biomedical industry received an indication last week that the National Institutes of Health (NIH) may reverse a policy known as 'reasonable pricing', under which the NIH can set a 'reasonable price' for products developed out of collaborative research projects.

The clause has been included by the NIH since 1989 in all Cooperative Research and Development Agreements (CRADAs) — introduced by Congress in 1986 to encourage technology transfer from federal laboratories — made with research groups in its intramural programme. The move followed controversy over the price of the anti-retroviral drug AZT.

But the clause is deeply unpopular with the biomedical industry, which sees it as a potential vehicle for imposing price controls on biomedical products, and argues that it is therefore a major barrier to collaborative projects with scientists working in the institutes' \$1.1-billion in-house programme.

The reasonable pricing clause was the hottest topic under debate at a meeting convened last week by NIH to discuss CRADAs. The panel included not only industrialists, academics and patient advocate groups, but also NIH officials.

Many industrialists feel that other provisions within the agreements already give the government adequate protection.

The industry is also worried that after spending money on research, trials and commercialization, the NIH, under congressional pressure, might exercise the reasonable pricing clause in its CRADA in a way that the industry considers unreasonable.

Chris Doherty, Washington director of a coalition of research hospitals, industry and patient advocate groups, says that, as a result, industry is not taking full advantage of the NIH's intramural programme. "This is a foolish situation," he says.

The biomedical industry's complaints about reasonable pricing have been fanned by uncertainty over the price controls that may accompany President Bill Clinton's proposed health-care reform.

Alison Taunton-Rigby, president of Mitotix, a biotechnology company in Cambridge, Massachusetts, says that venture capital companies have a checklist for companies they are considering investing in. "One of the things they ask is whether you have any CRADAs that include a reasonable pricing clause," she says. "It is a black mark if you do."

NIH officials acknowledge the problem. Publicly, they are not prepared to say that the reasonable pricing clause is likely to be dropped in all cases. But many feel that it should be.

Lisa Raines, vice-president for government affairs at the biotechnology company Genzyme, suggested an alternative clause insisting that a drug company should provide reasonable access to patients who cannot afford a particular therapy. Neither Mitotix or Genzyme are prepared to enter CRADAs with the NIH under the current terms.

Much of the discussion at last week's meeting centred on finding arguments for the panel's final report supporting a proposal that reasonable pricing should not be mandatory in all cases. Any change would take time to make its way through the federal bureaucracy. It could also run into opposition from Congress, some of whose members are keen that the principle embedded in the clause be maintained.

But industrialists at the meeting expressed optimism. "If we can eliminate this clause, we are prepared to sit through the months of negotiations that are a part of doing business with the government," said Taunton-Rigby.

Helen Gavaghan

Former KGB agent clears Bohr of passing secrets

Moscow. Recent claims by the former Soviet agent Pavel Sudoplatov that Niels Bohr was involved in passing US atomic bomb secrets to Soviet agents have been contradicted by the evidence of a former KGB scientist sent to obtain information from the eminent Danish physicist in 1946.

The accusations were included in Sudoplatov's book *Special Tasks*, published earlier this year, which has been widely criticized for its accusation that many leading Western physicists, in particular Robert Oppenheimer, helped the Soviet bomb programme by providing critical scientific information (see *Nature* **368**, 799; 1994).

But the diary of Yakov Terletsky, a physicist and a former employee of the KGB, the Soviet intelligence service, published in a



Bohr: 'gave nothing away'.

recent issue of the journal *The Issues of Science and Technology History*, reveals a different picture. It recounts how in 1946 Lavrenti Beria, chief of security for the Russian bomb project, ordered the KGB to contact Bohr, who at the time had a reputation for supporting open international relations in the area of atomic warfare.

Terletsky later told researchers that he asked questions prepared for him by Soviet scientists such as Igor Kurchatov, Yuly Khariton and Isaac Kikoin. "Bohr calmly answered all questions, but his answers were general," says Terletsky in the diaries.

Kurchatov subsequently reported to Beria that the 'Bohr interrogation' was of little value. He said that Bohr had not provided any new material, and that most of the information was already publicly available.

Terletsky said the operation had been largely a formality, designed to rebut charges that the KGB was not doing enough to help those working on the bomb.

According to Terletsky, his meeting with Bohr was intended primarily to indicate to Kurchatov that he could not put the blame for the delay on a lack of intelligence data by providing him with an opportunity to address questions to Bohr himself.

Vladimir Vizin, a specialist in the history of the atomic project, suggests that Kurchatov, Khariton, Artsimovich and Kikoin may have been reluctant to endanger Bohr through suggestions that he was a Soviet spy, and therefore compiled questions whose answers would contain little information.

Vladimir Pokrovsky

Silk worms fall prey to night stalker

New Delhi. A small lizard is being blamed for sabotaging an Indian research project to breed superior silkworms through genetic engineering. The centimetre-long lizard had been feasting nightly on silkworm eggs that had been genetically manipulated and left in a cage to hatch.

Elimination of the nocturnal invader has finally saved the research project. "But we nearly lost six months' work," says K. P. Gopinath, head of the microbiology department at the Indian Institute of Science in Bangalore.

Under a project funded by the Department of Biotechnology, Gopinath has been

working to improve the Indian mulberry silkworm (*Bombyx mori*) by introducing beneficial genes from a Chinese strain. The project was going well until staff noticed that the eggs and the larvae were disappearing from their cage. "One morning we would count 30 eggs, the next day it will drop to 15, and the following day the cage will be empty," said Gopinath. After three months staff decided to keep the cage under 24-hour surveillance, and spotted the lizard entering the cage through a small opening after the lights had been extinguished.

K. S. Jayaraman

Clinton's green think-tank under a cloud

Chicago. A new alliance between scientists and the private sector can solve some of America's most intractable environmental problems, Bruce Babbitt, the US interior secretary, told the President's Council on Sustainable Development last week when it visited the mid-West to see for itself how such problems are afflicting the country's heartland.

But the high-powered council, established last year by President Bill Clinton to help him to develop new and convincing environmental policies, was starkly confronted with the limitations of the Babbitt approach on the banks of Lake Michigan.

In particular, it faced the clear dichotomy between the shiny statements it has been polishing in Washington, and the dirty reality of life for those left to rot on what is known as the 'rustbelt'.

The Chicago meeting was the second such mission for the 25-strong council, made up of cabinet members, chief executives of leading industrial companies and the heads of the main environmental pressure groups. The council has drafted a 'vision statement' and 15 proposed principles, and is to report to the president in a year's time.

After field trips to Chicago's decrepit south side and to the steel town of Gary, Indiana, Jonathan Lash, co-chairman of the council and president of the World Resources Institute, said that the trips had stimulated "a very strong reaction" in council members. "A lot of people saw and heard things they hadn't seen before."

Lash says that the new experiences con-

cern "the whole question of brownfields and environmental equity". He was referring to the stalemate over the fate of disused industrial land that is adversely affecting people's lives, particularly those of the black people who live next door to it.

Gary was once the home of the world's largest integrated steel mill. The council's visit was intended to see local "ecological restoration activities", including the 20-year-old 15,000-acre Indiana Dunes National Lake Shore. This is the world's first national park on industrial land, and is a key site in the development of ecological science.

But the council heard of what Carlos Tolliver, a local activist, called "a city under economic and environmental siege". His description was hard to dispute at a public meeting held beneath the pall of thick black smoke from a local tyre dump, where 70,000 tonnes of rubber had been burning out of control for three days.

More landfill dumps are planned for Gary, where a quarter of households have an income of less than \$10,000 a year, and the city council is keen for any development it can get, sustainable or otherwise.

If there is a traditional US model for resource management, it is one in which environmentalists fight industrialists or landowners in the courts until an expensive stalemate is reached. Babbitt thinks that

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Great Lakes: test-bed for anti-pollution policies.

more thorough ecological science, together with greater cooperation from industrialists or landowners, can break the deadlock.

But even if his approach offers a way forward for resource-rich rural regions, it offers little for America's extensive areas of urban blight. One problem for the President's Council on Sustainable Development is that its tortuously wide brief covers both sets of problems.

Another is the recurrent allegation that the council is a cover for inaction on key areas of environmental policy. "We don't want fuzzy commissions — we want hard commitments," says Brett Hulsey of the Sierra Club, a powerful environmental organization, who complains that Congress is preparing legislation on clean water and wetlands conservation that is weaker than existing law.

Lash says that the council's brief is to do the kind of strategic thinking that government agencies cannot do on their own. Madeleine Kunin, the deputy education secretary, says: "The very fact that these groups are all collaborating for the first time gives us a chance to create a vision, and not just react to crises."

Some want the council to raise its profile. But environmentalists are enthusiastic about the council's endorsement of the so-called "precautionary principle" stating that "where public health may be adversely affected, or environmental damage may be serious or irreversible, prudent action is required even in the face of scientific uncertainty".

Sustainable development already means different things to different people. Asked what it meant to Chicago, Richard Daley, the city's mayor, said: "It's how you are going to deal with those wastelands." To Clinton, it means keeping faith with a sizeable constituency that voted for him partly or wholly on environmental grounds, without sacrificing jobs or growth. The council is to tell him how to reconcile these two goals. As Daley said: "It's a very, very difficult assignment, but people here appreciate your commitment to it."

Colin Macilwain

EPA pledges more long-term research

Chicago. The US Environmental Protection Agency (EPA) has announced plans to allocate half of its annual research budget to long-term research, to ensure that a greater proportion of its research funds is made available to universities and to improve standards of peer review.

In addition, the agency's laboratories will be reorganized into four sections dealing with risk management, risk assessment, exposure hazards, and health and environment issues.

The reforms were announced last week by Carol Browner, the EPA administrator, to a meeting of the President's Council on Sustainable Development in Chicago. At present, the bulk of the EPA's widely criticized research programme is short-term work needed to support the \$7.2-billion agency's regulatory function.

Browner hopes that the changes will head off calls for a free-standing National Institute for the Environment (NIE) aimed at raising the standards of federally funded environmental research.

The EPA was founded in the early 1970s as an amalgam of existing regulatory agencies, but has never established a significant track record in supporting environmental research. Its intramural programme is split up between a large number of laboratories, and 90 per cent of its \$350-million extramural programme is spent on routine contract research.

The reorganization of EPA's laboratories is intended to address the problems caused by fragmentation. But the largest changes will be those needed to achieve a 50:50 balance between short-term and long-term research; Neil Dobbs of the Campaign for an NIE claims the present balance is about 85:15 in favour of regulatory work.

The agency is promising to raise its annual funding for competitive, peer-reviewed, investigator-initiated grants from \$20 million to \$100 million annually, although no timescale for achieving this has been announced, nor was an EPA spokesman able to say where the money would come from.

Colin Macilwain

German academics find career boost in east

Munich. A survey carried out by *Nature* has confirmed that a large proportion of senior academic posts in higher education institutions in the former German Democratic Republic (GDR) have been given to west German academics.

The data were compiled from a survey of 57 east German institutions of higher education, including all major universities except the Humboldt University in Berlin and the Technical University of Dresden. Questionnaires were completed by 30 institutions.

The results reflect the success of efforts since reunification in 1990 to ensure that university posts are awarded to those best qualified to hold them. But they are also likely to fuel charges that west Germans are exploiting the situation in the east to further their own careers.

After reunification, the universities in the former eastern part of the country decided to adopt the west German higher education model, in which universities act as centres of both research and teaching, rather than being dedicated primarily to teaching.

As part of the change, all university teachers were checked for both their political and intellectual suitability to retain their posts. Those found to have actively collaborated with the Stasi, the former East Germany's secret police, or to be underqualified in western terms, were dismissed. Those remaining had to reapply for their posts through open competition.

Answers to the questionnaire show that in the politically neutral science subjects, the majority of the newly appointed academics are from the east; the proportion is two-thirds in science and technology, and

nearly three-quarters in medicine and veterinary science.

In the more ideologically sensitive humanities, by contrast, west German academics dominate. In the social sciences, nearly two-thirds now come from the west, and in economics around half. Only 10 per cent of the new law faculties are staffed by east German professors.

Although the Humboldt University did not take part in the survey, the latest general figures show that just over half of those appointed to the 387 positions offered so far came from west Germany and 90 per cent of the remainder were originally from the Humboldt.

Behind the aggregate figures lie some significant trends. Most of the movement from west to east involves young academics. Nearly half of the newly appointed professors under 50 years of age are from the west, compared with fewer than 20 per cent who are older than 50.

There is also a clear difference between the positions on the academic scale of academics under 50. Those on the highest rank (C4) are more likely to come from the west, while their eastern counterparts have had more success in competitions for lower rank positions.

And there is a clear difference between the types of higher education institutes. Whereas east Germans hold only 50 per cent of posts in universities, in the more technically orientated *Fachhochschulen* — institutes offering high-level vocational training — they hold 60 per cent of the posts. And east Germans have 70 per cent of the top-level professorships in *Fachhochschulen* compared to only 40 per cent in the universities.

The question of the number of west German academics who have taken jobs in the new *Länder* is politically very sensitive, and one on which many in the west are reluctant to comment.

But Gert Maibaum, head of the universities section of the research ministry of Saxony (the largest of the five new *Länder*), says that the survey results are not unexpected. Given time, he says, east Germans will be able to compete effectively with their western colleagues.

Josef Langer, general secretary of the University Rectors' Conference, agrees. "There is no disappointment among rectors," he says, referring to the relatively low success of east Germans in competing for

senior professorships.

There is, however, considerable disappointment among east German academics. Andre Hahn, education policy spokesman for the Party of Democratic Socialism (PDS) — the successor to the GDR's communist party — says that they were automatically at a disadvantage compared to their west German colleagues. For example, he says, the isolation of their community under communism meant that academics had less opportunity to publish their work in the scientific literature. Yet a major criterion for professorships was an applicant's publication record.

There is the serious danger that new professors coming from the west will stay in the east only for as long as it benefits their careers, and will leave as soon as a better opportunity arises in the west.

● A similar outcome was found by *Nature's* survey of non-university research establishments last year (*Nature* 362, 685; 1993) where most of the top jobs were snapped up by west Germans. In 1990, the science council (the Wissenschaftsrat) expressed the hope that there would be a "seeding" of these institutes with 10 per cent west Germans at all levels. But in fact few westerners took up lower positions.

Antje Langer

German science funds hold out against cuts

Munich. Paul Krüger, the German research minister, has persuaded his cabinet colleagues to approve a rise of 2.7 per cent in next year's research budget. If, as expected, this is endorsed by parliament, German research will have DM9.47 billion (US\$5.97 billion) of federal funds to distribute in 1995, DM250 million more than this year.

With an inflation rate of around 4 per cent, this still means a cutback in real terms. But it also means that research has done considerably better than other areas of public spending; next year's general budget will be kept down to a one per cent increase on this year.

Krüger says that he still intends to prioritize key technologies such as information and materials sciences and biotechnology, with increases of around 3 per cent and 9 per cent respectively. Environmental technologies will get an even larger increase of 10 per cent.

But basic science will not lose out. The research ministry will stick to its promise in 1990 to raise the Max Planck Society budget by 5 per cent for at least five years. Although only slightly higher than inflation, this agreement no longer appears as restrictive as it did a few years ago.

Alison Abbott

Statistical breakdown of higher education staff in east Germany

Origin by faculty	East	West	Other*
Medicine and Vet. science	78.0	19.8	2.2
Science and Technology	64.2	24.7	11.1
Social sciences	21.3	61.3	17.4
Humanities	34.2	42.9	22.9
Economics	31.9	45.6	22.5

Age 30–50

Universities	29.5	47.4	23.1
Fachhochschulen	42.2	53.0	4.8
C2 professors	56.1	36.2	7.7
C3 professors	39.1	47.7	13.1
C4 professors	21.6	55.8	22.6

Age 50+

Universities	69.1	17.2	13.7
Fachhochschulen	78.1	18.7	3.2
C2 professors	89.8	4.3	5.9
C3 professors	77.7	15.3	7.0
C4 professors	64.2	21.5	14.2

All figures expressed as percentages

* 'Other' includes staff employed on short term contracts and positions filled by non-Germans

Switzerland balks at opening research funds

Basel & Munich. Switzerland has been told that it can participate in the European Union (EU) fourth Framework research programme provided that it agrees on a package of political demands from the union — and opens up some of its own national research programmes to scientists in EU states.

But the EU research ministers have held back from offering the same deal to Israel on the grounds that, unlike Switzerland, it lies outside the European continent, and is therefore not eligible for EU membership.

The ECU12 billion (US\$13.6 billion) Framework programme, which runs from 1995 to 1999, is designed partly to stimulate European competitiveness by supporting research that underpins the future needs of

But science administrators in Switzerland are worried about a demand from the EU research ministers that, in return — and to ensure a reciprocal exchange of scientific results between the EU and Switzerland — the country should open up its own 'equivalent' national programmes to grant applications from scientists in other EU countries.

A spokesman for the Swiss foreign ministry points out that the national science funding agencies in EU member states are not required to do this, and describes the demand as "surprising". Hans Peter Hertig, general secretary of the Swiss National Science Foundation (SNF), goes further, and says that the condition is "unacceptable".

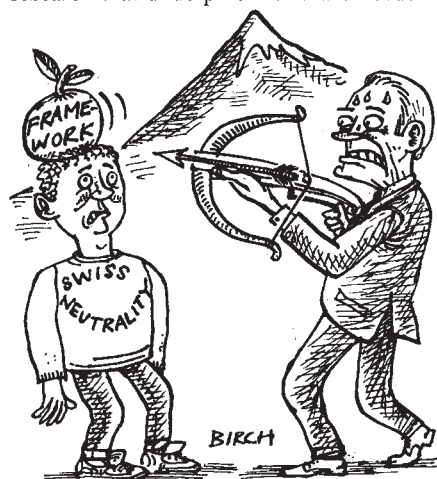
In response to such concerns, commission officials concede that it may be necessary to negotiate a compromise. One possible solution could be to identify specific research programmes, perhaps part-

funded by industry, that would be opened up to wider participation. The SNF, which supports basic research, would continue to fund Swiss researchers only.

Israel's request for similar full participation in the Framework programme was rejected by the council of research ministers at a meeting last month because some EU countries, in particular France, feared that such participation would set an undesirable precedent for other non-European countries.

But Israel, backed by the commission itself, is still hoping to persuade the ministers to change their minds. Marcel Shaton, from the Israeli embassy in Brussels, says that the high standard of science in Israel means that participation would be of mutual benefit. He also says that it would help to promote good industrial relations between Israel and Europe.

Oliver Klaffke & Alison Abbott



industry. Some EU countries argue that this goal could be endangered if the programme is opened up to non-European states.

The Swiss population has consistently rejected all suggestions that Switzerland should join the EU. But the government, reflecting the view of its industrial community, is still keen that it should participate in Framework, and has made such a request to Brussels.

A final decision will depend on the outcome of separate negotiations over the next six months on Switzerland's adherence to certain EU principles, including freedom of commercial traffic through the Alps (which Switzerland has voted to confine to rail within 10 years), and the freedom of EU citizens to live and work in Switzerland.

If these negotiations are successful, and Switzerland is subsequently accepted into Framework, it will have to contribute ECU375 million to the programme over five years. Bruno Spinner, a member of the Swiss delegation responsible for EU negotiations, says the money will be well spent. "A considerable amount of money will flow back into Switzerland as a result of participation," he adds, in a bid to reassure critics who argue that the contribution will be taken out of the science budget (see *Nature* 369, 264; 1994).

UK bans use of fetal eggs in IVF

London. Britain's Human Fertilization and Embryology Authority (HFEA) announced last week that it will not allow immature oocytes or ovarian tissue from aborted fetuses or cadavers to be used in the treatment of infertile women.

But it has agreed that such tissue can be used for research. It also said that it is prepared to approve the use of ovarian tissue from adult cadavers for treatment "in principle", but will do so in practice only when wider issues about consent have been investigated further.

The HFEA's decision comes after a six-month consultation period during which the body, set up by the government almost four years ago to regulate both research on and the application of fertilization treatment, received nearly 10,000 responses.

Sir Colin Campbell, chairman of the HFEA, said last week that its main concerns about the use of fetal ovarian tissue for treatment were the possible effects on the recipient mother, the strength of public opposition (83.2 per cent of respondents were against it), the unknown psychological effects on a child born in this way, and lack of knowledge of the development of fetal eggs.

But Campbell added that he was optimistic that the need for the technology, which is still 10–20 years away from being feasible in humans, would be reduced or eliminated by the development of other methods of fertilization. The two areas considered most promising are to take pieces of ovarian tissue from volunteers and mature the oocytes *in vitro*, or to transplant the donated ovarian tissue into infertile women where the eggs could mature naturally.

Research using fetal ovarian tissue will

be licensed by the authority, subject to existing controls, including the need for informed consent to the use of the tissue for research.

Almost 60 per cent of respondents were opposed to this. But Campbell said that research could do much to alleviate disease, improve processes for fertility treatment and prevent the passing on of inherited diseases, and because "ultimately you can't stop research". He added later that even if public opposition meant that this research was banned in the United Kingdom, it would be carried out elsewhere in the world.

The HFEA's conclusions effectively coincide with an amendment pushed through the House of Commons in April by Dame Jill Knight, Conservative member of parliament for Birmingham-Edgbaston, which stated that fetal germ cells should not be used "for the purposes of providing fertility services for any woman".

A second attempt by Lord Walton to overturn the ban had failed in the House of Lords the previous week, after help previously offered by the government to reword an earlier amendment in a more acceptable form had failed to materialize.

Commenting on the parliamentary moves, Campbell said, in the context of the HFEA's advice, Knight's amendment was "quite redundant".

The legislation does not differ from the HFEA's recommendations in what it prohibits. But Richard Gardner of Oxford University, who chaired a working group which prepared a response to the HFEA on behalf of the Royal Society, criticizes the fact that Knight introduced her amendment before the consultation process had run its course.

Maggie Verrall

Plans for US wind tunnel set to take off

Washington. The US National Aeronautics and Space Administration (NASA) seems likely to win funding for the construction of the world's largest and most sophisticated wind-tunnel complex, to be used by aircraft manufacturer Boeing to speed up the development of successors to the Boeing 747 jumbo jet.

A proposal earlier this month by a Senate budget subcommittee to put \$400 million into the \$2.5-billion project in the next financial year, beginning in October (see *Nature* 370, 167; 1994) may be scaled back. But congressional staff say, with broad support within the Clinton administration and both houses of Congress, it will almost certainly open the way for funding to commence in the following year.

The project will involve the construction of huge subsonic and transonic wind tunnels, and is significant because it represents direct support to the US civil aerospace industry. Until now, such support has tended to be indirect, through military and space contracts for example.

A number of industrial companies will use the tunnel. And universities say that the capabilities of the tunnel complex will make it a potentially valuable tool for basic research in aeronautics.

But the facility, to be known as the National Wind Tunnel Complex, will be designed primarily for the convenience of air-

craft development engineers wanting to test and refine aircraft prototypes quickly and methodically. In particular, its main purpose will be to help Boeing to compete with the European Airbus consortium, its only

nautical engineering department of the University of Southern California, Los Angeles, existing wind tunnels range up to one million, whereas large modern airliners can have Reynolds numbers of 100 million.

The proposed NASA tunnels will be large, but not excessively so — the larger of the two will have a 24 by 21 foot cross-section. Their high price reflects the power and structural strength needed to blow air through them at sufficient speed and pressure, with a minimum of vibration or background turbulence.

Congressional supporters of the tunnel are concerned that US aircraft manufacturers currently rely on European facilities, especially the German-Dutch wind tunnel at Emmeloord in the Netherlands, for their development work, and are keen that the United States should regain its lead.

"The argument is that the complex will position the US industry to prevail, not merely achieve equity [with Europe]," says Larry Ross, director of NASA's recently established Wind Tunnel Program Office. "The facility would be better than anything else in the world."

The prospect of a new \$2.5-billion high-technology project has set off a scramble by congressmen and senators wanting it located in their home state. Political and logistical considerations both point strongly to the West Coast.

Colin Macilwain

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Europe's German-Dutch wind tunnel: US wants to do better.

rival in the market for large civil aircraft.

The need for the two tunnels was identified in a study carried out last year by NASA, the Department of Defense and other federal agencies to identify important areas of deficiency and overlap in US research facilities. It found no need for extra supersonic or hypersonic wind tunnels, so the proposed complex will be for the development of conventional, subsonic airliners, rather than the supersonic ones which NASA is also keen to support.

The ability to derive useful information on airflow from scale models in a wind tunnel depends on a dimensionless parameter called the Reynolds number.

According to Hsien Cheng, of the aero-

Soros fund offers helping hand to European agency

Moscow. The International Sciences Foundation (ISF), set up by the Hungarian-born financier George Soros to support scientists in states belonging to the former Soviet Union (FSU), has offered to help the parallel body created by European governments to distribute its grant money in Russia and avoid paying heavy state taxes.

The offer has been made to INTAS — the International Association for the Promotion of Co-operation with Scientists from the Independent States of the FSU — which was established by the member states of the European Union (plus Austria) in 1992.

ISF officials say they made the offer in recognition of the difficulties INTAS has faced in ensuring that funds reach the scientists for whom they are intended through official channels, without having either customs duties or income tax deducted.

They claim that many Russian scientists and colleagues in the West have become frustrated at the delays caused by these difficulties, and have offered the ISF's formal exemption from taxes, negotiated per-

sonally by Soros with the Russian president Boris Yeltsin, as a way around the problem.

But INTAS officials have rejected the ISF offer. They claim that, in principle, the charitable status of their organization means that its grants are automatically exempt from taxes, and that recent negotiations with European and Russian banks have led to a new agreement on procedures for channelling funds directly to Russian scientists.

They also argue that their decision is a result of their more decentralized mode of operation. "It is our policy not to centralize things too much," says Pierre Vennay, the European organization's director-general.

Russian legislation requires more than half of any grants sent into the country to be paid to the government in taxes and customs duties. INTAS member states refused to transfer funds under such conditions. "Our organization is not intended to support the Russian state," says Flemming Woldbye, chairman of the INTAS general assembly.

The difficulty has been compounded by

doubts among scientists in both Russia and the West that money sent through formal channels will indeed arrive in the pockets of scientists untaxed, despite official reassurances to the contrary.

Vladimir Skulachev, chairman of the ISF Russian advisory council, says that making use of the ISF would enable INTAS to cut down on organizational costs, as well as to use the bulk discounts that the ISF has negotiated on the purchase of scientific equipment in the West.

But INTAS remains optimistic that it can solve the problem on its own. Vennay acknowledges that setting up the channels to transfer funds has been a lengthy and complex task, but says that money is now being transferred to Russian scientists on an experimental basis.

Furthermore, Woldbye says that at a recent meeting with INTAS managers, Boris Saltykov, the Russian Minister of Science, promised to discuss tax and customs duties with his colleagues in the Russian government.

Vladimir Pokrovsky

Japanese centre set to boost efforts in brain research

Wako City, Japan. The Institute of Physical and Chemical Research (RIKEN) in Wako City in the northwest suburbs of Tokyo is set to become one of the most powerful centres for brain research in Japan.

In October the brain research section of the institute's highly successful Frontier Research Programme will double in size through the opening of three new laboratories, and further expansion is already being planned.

RIKEN's Frontier Research Programme has become widely known since it was established in 1986, partly because its research projects are closely monitored by external reviewers, and partly because of deliberate efforts to draw on the talents of foreign as well as Japanese scientists. Both are unusual for Japan.

The programme started with two broad research fields, biohomeostasis (which includes research on areas such as ageing and intestinal flora) and 'frontier materials', which initially focused on quantum devices, nonlinear optics and bioelectronic devices. Research on brain mechanisms of mind and behaviour was added in 1988.

In addition, two Frontier research centres have been opened on other sites. One in Sendai, north of Tokyo, opened in 1990 and focuses on photodynamics research. The other in Nagoya, southwest of Tokyo, opened last year and specializes in research on biomimetic control: it aims at combining RIKEN's expertise in brain and materials research and in the development of sensor technology with robotics research and studies of the molecular biology of the brain.

Three new laboratories for research on brain information processing will be added in October at RIKEN's Wako City campus. These are the Laboratory for Neural Modeling, the Laboratory for Information

Representation and the Laboratory for Artificial Brain Systems.

RIKEN officials say they hope to win approval for several more brain research laboratories in next year's budget. A decision on whether to apply to the Ministry of Finance for funding will be made by the Science and Technology Agency, which provides most of RIKEN's budget, at the end of August.

Brain researchers are "scattered in many small groups" in universities around Japan, says Katsutoshi Yasufuku, an administrator of the Frontier programme. He says RIKEN means to bring together a large and powerful group at a time when there are several leading researchers in neuroscience in Japan.

In cooperation with the new brain research groups, the institute also hopes to develop brain scanning technology, such as magnetic resonance imaging, positron emission tomography, and scanners for detecting minute magnetic fields in the brain, as well as techniques for artificially growing and keeping brain slices.

Neuroscientists at RIKEN are led by Masao Ito, one of Japan's leading brain researchers and also the director-general of the whole Wako Frontier programme. They are so confident of the quality of their research that they are already planning to request a second external review of the brain research laboratories during their first phase of operation.

External reviews take place at two levels. The Wako, Sendai and Nagoya centres each have an advisory council made up of leading Japanese and foreign scientists. The councils, whose members change every three years, meet once or twice a year to review the research and suggest future directions for the programme at each site as a whole.

David Swinbanks

Japan edges cautiously towards gene therapy

Tokyo. Japan took another cautious step towards its first application of gene therapy last week when a medical ethics committee at Hokkaido University in Japan's northernmost island gave "emergency" approval for the genetic treatment of enzyme deficiency in a 3-year-old boy.

The proposed treatment is for adenosine deaminase (ADA) deficiency, which can lead to a breakdown in the immune system, and the approval has been requested by Yukio Sakiyama in the university's department of paediatrics.

The proposal was submitted to the university committee last November. But

it was not approved until last week because of the tardiness of the Ministry of Education, Science and Culture, which has administrative jurisdiction over all Japan's universities, in introducing a system for screening such experiments.

More bureaucracy lies ahead. Both the Ministry of Health and Welfare and the education ministry have set up committees to screen such experiments, and approval is required from both — even though some scientists belong to both committees (see *Nature* 369, 5; 1994). Sakiyama says he hopes to submit "basically identical" proposals simultaneously to the two committees in the near future. □

French ministers face public trial in HIV blood affair

Paris. The French dispute over responsibility for the distribution to haemophiliacs of blood contaminated with human immunodeficiency virus (HIV) in the mid-1980s erupted again last week with a decision to permit charges to be heard against three former socialist ministers who were in office at the time.

The decision results from a reform of the High Court of Justice last year which allows any individual — and not just parliament — to bring charges against ministers, in this case Laurent Fabius, the then prime minister, Georgina Dufoix, then minister of social affairs, and Edmond Hervé, then deputy health minister (see *Nature* 364, 269; 1993).

The Court of Justice of the Republic last week accepted 11 civil suits against the three politicians. Indeed the reform was itself prompted by the failure of repeated attempts to try the three politicians.

These will probably now face criminal charges of deliberately administering substances dangerous to health. If convicted they could face prison sentences of five to ten years. The charges would cover the period between March 1985 — the date when the government is acknowledged to have been informed that unheated blood-clotting factors were contaminated with HIV — and October 1985, when the French National Blood Transfusion Service (CNTS) stopped their distribution.

The prospect of charges against the three politicians is likely to give renewed prominence to the affair, particularly as Fabius is a potential candidate in next year's presidential election.

Attention will focus on the role played by the ministers' scientific advisers — against whom separate charges are pending — in the delays in introducing both heat-inactivated blood products and routine screening of donated blood for HIV (see *Nature* 367, 304 & 673; 1994).

In a separate development, it was announced last week that Michel Garretta, former director-general of the CNTS, is to be retried on the criminal charge of 'poisoning'. Garretta had already been sentenced to four years' imprisonment in 1992 on the misdemeanour of "deception over the quality of products sold".

The decision to press further charges against Garretta follows a ruling last month by the Paris Supreme Court of Appeals.

Garretta's lawyers have already said that they will fight the new decision in the European Court of Human Rights, pointing out that the European Convention on Human Rights states that no one can be tried on different charges for the same act.

Declan Butler

Regulatory disincentives

SIR — The report by Bragg *et al.* describing the effectiveness of bioremediation for the 1989 Exxon Valdez oil spill in Prince William Sound, Alaska, is both cogent and potentially useful¹. However, the actual bioremediation techniques employed represent science and technology worthy of the nineteenth century. Many biological scientists, particularly those within the 'biotechnology' community, must echo the sentiments of Mr William Reilly, who was the administrator of the US Environmental Protection Agency (EPA) at the time of the incident: "when I saw the full scale of the disaster in Prince William Sound in Alaska . . . my first thought was: where are the exotic new technologies, the products of genetic engineering, that can help us clean this up?"²

For the most part, innovative products of the new biotechnology for bioremediation remain on the drawing board, with researchers and companies intimidated by regulatory barriers and disincentives. In the United States, for example, the EPA has tried unsuccessfully for a decade to articulate a final policy on field trials with microorganisms — including those for bioremediation — but has returned again and again to proposals that discriminate against microorganisms created with high-precision rDNA technology while exempting research organisms crafted with any other technique. Indeed, because it is anticipated that, for the foreseeable future, regulatory submissions for bioremediation research on rDNA-manipulated microorganisms will require "excessive time, money and paperwork", and because of the uncertainty of ever obtaining a licence, the US bioremediation industry has largely restricted itself to work with naturally occurring organisms². But because of the nature and complexity of the substances involved in most spills and toxic wastes, naturally occurring organisms seem not to be up to the job.

The regulatory climate is no better outside the United States. For example, the European Union (EU)'s regulatory approach was taken to task as regressive and anti-innovative in the October 1993 report of the UK House of Lords Select Committee on Science and Technology, *Regulation of the United Kingdom Biotechnology Industry and Global Competitiveness*³. Echoing other assessments of the EU's environmental biotechnology policies, the report recommended reduced and rationalized regulation, concluding that: "As a matter of principle, GMO [genetically manipulated organism]-derived products should be regulated according to the same criteria as any other product . . . U.K. regulation of the new biotechnology (according to EU directives) of genetic modification is ex-

cessively precautionary, obsolescent, and unscientific. The resulting bureaucracy, cost, and delay impose an unnecessary burden to academic researchers and industry alike."³

Regulatory disincentives are potent. Until governments demonstrate rationality and sensitivity to scientific principles in their regulatory policies toward the new biotechnology, we'll be slopping bacterial growth media on the beach instead of putting high technology to work.

Henry I. Miller

*Hoover Institution Institute
for International Studies,
Stanford University,
Stanford, California 94305-6010, USA*

1. Bragg, J. R., Prince, R. C., Harner, E. J. & Atlas, R. M. *Nature* **368**, 413–418 (1994).
2. Day, S. M. *Trends Biotech.* **11**, 324–28, (1993).
3. Ward, M. *Bio/Technology* **11**, 1213 (1993).

Health research

SIR — I should like to clarify a point in David Dickson's report about research supported by the UK National Health Service (*Nature* **369**, 514; 1994). The teaching element of the Service Increment for Teaching and Research (SIFTR) related to student numbers generally "works well", although the sums involved may be insufficient. The Council of Deans agreed recently, however, that the R element should be separated from the T and be allocated in relation to research quantity and quality. The deans' concern was that the R element, designed to support the research infrastructure of university hospitals essential to a wide range of clinical and health services research, might be transferred either to purchasers of health services or to the currently insufficiently funded Health Services Research programme, which would 'rob Peter to pay Paul' and greatly damage the research capability of university hospitals.

Peter Richards

*(Dean)
St Mary's Hospital Medical School,
Norfolk Place, London W2 1PG, UK*

Re-setting the great count

SIR — Emiliani¹ proposed to solve the problems with our calendar by starting a new count at the beginning of the Holocene, specifically by establishing the birth of Christ as the year 10,000. But one might just as well set the death of Buddha or the Hegira of Mohammed at the year 10,000 and gain other valid and artificially exact starts to the Holocene. Emiliani

therefore fails to address one concern he raised himself: that marking time using the event of Christ's birth has no significance for many of the world's cultures.

I have a different solution². No past event (the birth of Christ, the death of Buddha, the Woodstock Rock Festival) can fill the requirement of being truly global. A new calendar must await, therefore, some future moment, as only recently have the world's peoples been sufficiently linked to have such an experience. What could it be? A great war? An asteroid impact? I propose a more positive possibility: When the world is ready it could agree to re-set the count (for example, through a vote in the United Nations). This itself would be of sufficient historic importance to kick off the new year number one — the united agreement to start would be the event implicitly celebrated ever after by the new calendar.

Tyler Volk

*Earth Systems Group,
New York University,
34 Stuyvesant Street,
New York, NY 10003-7599, USA*

1. Emiliani, C. *Nature* **366**, 716 (1993).
2. Volk, T. *Metapatterns across Space, Time, and Mind* (Columbia University Press, in the press).

Out of tune

SIR — Writing from San Diego (*Nature* **369**, 270; 1994), Lawrence Bruton calls for a more logical system of nomenclature in membrane biology. If his proposed scheme is adopted, he concludes, 'dotriacontahectaspan' will roll off the tongue as readily as 'hemisemidemiquaver'. But the correct word in the United Kingdom is 'hemidemisemiquaver', and the United States already has a more logical system of nomenclature in music, giving 'sixtyfourth-note'. For biomusicians this casts doubt on all that went before. . . .

Peter B. Soul

*51 Lakeside,
Earley, Reading, Berks RG6 2PG, UK*

Long on authors

SIR — Writing on "making [scientific] publications more respectable" (*Nature* **369**, 353; 1994), you suggested that every co-author of a paper should, in principle at least, be able to give a brief public talk on the substance thereof. The first article of this same issue of *Nature* (2 June), if I counted correctly, had 108 authors. The public forum that puts this suggestion to the test is going to be a lengthy one.

Beverly Griffin

*Department of Virology,
Royal Postgraduate Medical School,
Du Cane Road,
London, W12 0NN, UK*

COMET SHOEMAKER-LEVY 9

Dazzling demise of a comet

Clark R. Chapman

WHAT happens when a comet plunges into a planet? After a year of mounting speculation, last week we finally found out. Periodic comet Shoemaker-Levy 9, its 20-odd 'pearls' by then strung out over more than 300 million kilometres, has met its end; and its target, the gas giant Jupiter, ten times the radius of the Earth and more massive than the rest of the planets put together, bears the scars.

As the crash approached, comet researchers hesitated to predict more than a modest display — in Paul Weissman's phrase¹, a 'cosmic fizzle'. Perhaps they recalled the underwhelming reality of 'comet-of-the-century' Kohoutek in the early 1970s. But nature had something else in store for the worldwide, electronically connected network of astronomers who awaited the first impacts on Saturday evening, 16 July. The spectacular show during the subsequent week exceeded even the most optimistic predictions about the events' potential visibility from the Earth, nearly 800 million kilometres away. No one doubted that significant events would occur at the impact sites around Jupiter's limb, hidden from our sight. But the prominent visibility of high plumes, hotspots and enormous, long-lasting dark splotches on the face of Jupiter was wholly unexpected. What does it mean?

On the practical side, it has meant a revival of interest in amateur astronomy. Even the United States House of Representatives got into the act last week, adding language to NASA's authorization bill requiring the space agency to search for all the asteroids and comets that threaten the Earth. But the longer-term benefits to astronomy will come from years of analysis of the unprecedented wealth of information from this unique phenomenon.

As results continue to pour in over the Internet from all corners of the world it is difficult to synthesize all the observations. There seem to be contradictory implications from some of the different sets of data. But evidently none of the pre-crash models is entirely correct, which is not surprising for such an unprecedented event.

One of the most impressive and unexpected aspects of the comet impacts was the dramatic plumes that reached more than 2,000 km above Jupiter's cloud tops, clearly depicted in pictures by the Hubble Space Telescope. Some numerical mod-

ellers had tentatively hoped that one or two of the later impacts, which occurred much closer to the Earth-facing side of Jupiter, might manage to peek above Jupiter's limb before thinning into invisibility. Instead, the plumes either were propelled to higher altitudes or remained much more opaque (and hence capable of



The collision of fragment G, pictured 12 minutes after impact. The impact site of fragment A can still be seen dimly on the opposite limb. Image at 2.34 μm with CASPIR by Peter McGregor, using the ANU 2.3-m telescope at Siding Spring, Australia on 18 July.

reflecting sunlight) than had been expected. Consequently, pictures of Jupiter taken through filters in deep methane absorption bands near 2 micrometres (where Jupiter is very dark) show dramatically bright flares shortly after impact. These are the plume tops that have ascended above Jupiter's atmosphere and catch the sunlight, undimmed by methane absorption. They remained bright at 2 μm for hours and days after the plumes fell back, persisting at altitudes where Jupiter's atmospheric pressure is only 2 thousandths the sea-level pressure on the Earth. Images taken at 10 μm (thermal infrared) show that the plumes were warm, and there is evidence that temperatures reached several thousand degrees.

Before the crash, there had been much debate about the size of the comet fragments. Were they large fragments, several kilometres in diameter, from a tidally disrupted body initially nearly 10 km in size, as calculated by Z. Sekanina and his colleagues²? (That was a view that comet co-discoverer Eugene Shoemaker continued to maintain even after Harold Weaver, analysing Space Telescope photographs of the comet, changed from

advocacy of large fragments to neutrality³.) Was the original body much smaller, with individual fragments only half a kilometre across, or smaller, as calculated Jim Scotti and Jay Melosh⁴? Were the fragments actually rubble piles or swarms of much smaller components, as predicted by Stuart Weidenschilling⁵, and modelled by Willy Benz and Erik Asphaug⁶? Or were the apparent fragments only ethereal wisps of dust, as some researchers privately began to fear after Hubble pictures last spring showed some of them to be vanishing like the Cheshire Cat?

A related question was how deep the comet pieces would plunge into Jupiter's atmosphere before exploding. If the fragments were large and cohesive, it was thought that they might dive more deeply, creating a buoyant fireball that would dredge up the water-rich lower-jovian atmosphere and spread it into the spectroscopically measurable stratosphere of Jupiter. Deep impacts might also induce seismic waves to spread through Jupiter's interior and re-emerge at the cloud surface, possibly with visible manifestations. Comet swarms, on the other hand, might yield upper-atmosphere meteor storms with no deep penetration.

It is widely assumed that last week's spectacular cosmic show means that the deep-plunging, large-fragment model is correct. But that is not necessarily so. For one thing, the penetration predic-

tions varied more because of different assumptions in the physics involved than because of differences in the assumed size and nature of the projectiles. And the swarm proponents never ruled out the possibility that a meteor storm might be as impressive as, or more so than, single explosions.

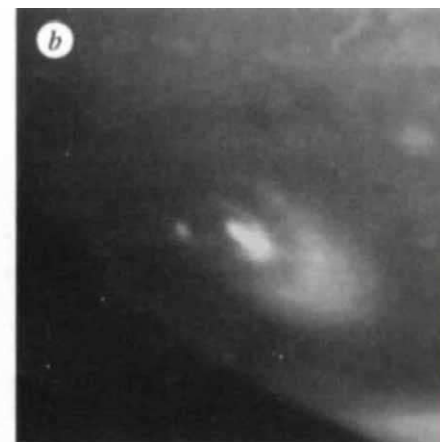
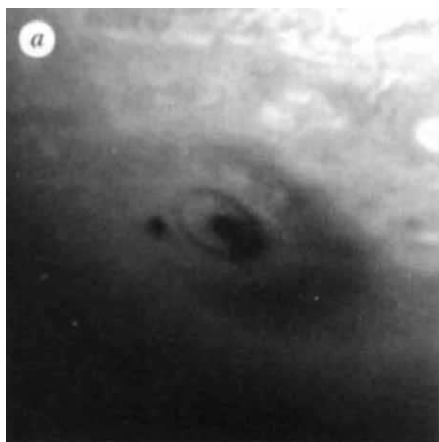
Let it be thought that all predictions were too pessimistic, consider large fragment K which struck on 19 July. It was first detected, at 2.3 μm , by Peter McGregor and Mark Allen at the Siding Spring Observatory in Australia. The fireball plume grew to be very bright. Yet they, and other observers in Australia, searched for and failed to detect any precursor brightening of Jupiter's satellite Europa. At the time, Europa was in sight of Earth and of the impact site, but hidden in Jupiter's shadow from any sunlight. Its shadowed, reflective, icy surface should have served as an excellent mirror for the expected brilliant meteor flash and the subsequent glowing fireball that would rush upwards from the explosion. Some researchers had predicted that this single opportunity for observing the reflection from an eclipsed moon of Jupiter would be



The succession of impacts leaves a glowing chain around Jupiter. The image here is taken in the 1.7- μ m band of methane emission using MAGIC on the 3.5-m telescope at Calar Alto, Spain on 19 July.

one of the most sensitive ways to detect an impact from Earth, even if everything else failed. But it turned out to be the other way round.

The spacecraft Galileo is best situated to resolve questions about optical flashes. Its first tape-recorded images are due to be played back to Earth on about 15 August and will show the impacts directly. The camera is very sensitive, so it should have recorded both the meteor flashes and any subsequent fireballs for six of the events even if the phenomena were much fainter than expected. There are already indications from Galileo's photopolarimeter-radiometer, which recorded two impacts, that the meteor flashes lasted longer than expected (about half a minute, at about 5 per cent of the total brightness of Jupiter), but that there were no fireballs. It may be back to the drawing board for the numerical modellers.



Aftermath of the G impact, as seen by the Wide Field Planetary Camera 2 on the Hubble Space Telescope shortly after impact. *a*, Taken with green light, *b*, taken with a methane filter. (H. Hammel, MIT/NASA HST.)

It appears, from preliminary spectroscopic data, that the fragments did not penetrate very deeply. The idea was to look for changes in very sensitive spectral lines, which might indicate contamination of Jupiter's stratosphere by material erupted from below. Evidently, little of Jupiter's presumed water-rich lower atmosphere was splashed up into the stratosphere. Most of the spectral changes can be explained by contamination from the dispersed comet itself. Yet other compounds such as hydrogen sulphide have been detected, which has never before been seen on Jupiter and which some researchers think is more likely to be derived from NH_4SH clouds below Jupiter's ammonia cloud deck

than from the comet.

The size of the comet before it broke up must have been large. Whether the main fragments consisted of monolithic objects, or clusters of innumerable smaller objects, the total mass that struck Jupiter must have been enormous to generate such high plumes and extensive dark patches on Jupiter's cloudy surface. Indeed the effects are so large, even for the largest estimated sizes of the comet fragments, that it seems likely that our understanding of atmospheric impact physics will have to be revised to account for the amazing phenomena that have been seen — as well as for what has not been seen.

Somewhat overlooked in all the excitement about the first impact by fragment A, and some of the impressive subsequent ones, is the fact that several impactors, including B, F, U and V, really did fizzle. Although photographs of the comet

earlier this year showed B to be as bright as A or C, there were only a couple of marginal reports of its impact plume, and it left only an insignificant black spot in Jupiter's atmosphere. Yet the Caltech Submillimetre Observatory seems to have detected hydrogen sulphide associated with B. Clearly, not all fragments of Shoemaker-Levy 9 were the same. One clue might be that some of the fizzled impactors had drifted off the main axis of the 'comet train' during the spring, like some of the fragments that prematurely vanished.

Unquestionably, the most dramatic evidence of the comet's demise is the belt of immense black patches left in Jupiter's high atmosphere. The larger ones are roughly the size of the largest spot on Jupiter — the Great Red Spot — and are much darker and more prominent than it in a small telescope. Even untrained observers using small, back-yard telescopes could easily watch the remnant of fragment G during the evenings following the impact. And most of the other fragments left black scars that were more visible than any of Jupiter's other usual spots. As these high atmospheric features are not grounded in the dynamics of Jupiter's lower atmosphere, like the Red Spot, it has been expected that the spots will fade and disappear over the days and weeks ahead. But they seem to be persisting. Because they are so black, and absorb so much sunlight, they could conceivably affect Jupiter's local atmospheric temperatures enough to generate longer-lived storms.

As some of the new black spots in Jupiter's southern latitudes are considerably larger than the whole planet Earth, sceptics about the potential effects of cometary impacts on Earth have fallen silent. News media and scientists alike started forgetting to add a caveat when they spoke of the link between the Chicxulub crater in Mexico and the extinction of dinosaurs 65 million years ago. Nothing about last week's impacts changes what is known about the incomplete fossil record of dinosaurs, but the Jupiter crash has certainly affected the way people feel about the role of impacts on our own planet. The black patches on Jupiter have converted planetary impact processes from the realm of theoretical possibilities and ancient geological history to a manifest, dynamic reality. □

Clark R. Chapman is at the Planetary Science Institute, 620 North Sixth Avenue, Tucson, Arizona 85705, USA.

1. Weissman, P. R. *Nature* **370**, 94–95 (1994).
2. Sekanina, Z., Chodas, P. W. & Yeomans, D. K. *Astr. Astrophys.* (in the press).
3. Weaver, H. A. *IAU Circ.* No. 5947 (1994).
4. Scotti, J. V. & Melosh, H. J. *Nature* **365**, 733–735 (1993).
5. Weidenschilling, S. J. *Nature* **368**, 721–723 (1994).
6. Asphaug, E. & Benz, W. *Nature* **370**, 120–124 (1994).

A glycation connection

Charles R. Harrington and Camilo A. L. S. Colaco

ALZHEIMER'S disease is a heterogeneous disorder characterized by disruption of selective neuronal cortico-cortical connections. The neuropathological hallmarks of the disease are neurofibrillary tangles and senile plaques, the main constituents of which are tau protein and amyloid- β ($A\beta$) protein, respectively. Much is known about the abnormal processing of tau and the precursor for $A\beta$ (ref. 1), but the mechanisms by which the insoluble protein deposits accumulate in the Alzheimer brain remain unclear. Two reports^{2,3} now provide evidence that glycation may be involved in the development of Alzheimer-type pathology.

Protein glycation starts with the non-enzymatic, spontaneous condensation of a sugar aldehyde or ketone group with free amino groups, such as ϵ -NH₂ groups of lysine, to form a labile Schiff base, a reaction first described by Louis-Camille Maillard in 1912. The resulting adducts then undergo irreversible rearrangements, to form more stable Amadori or Heyns products and the irreversibly crosslinked advanced glycosylation end-products (AGEs; see figure)⁴. These products are a heterogeneous group and result from a complex cascade of dehydration, condensation, fragmentation, oxidation and cyclization reactions. Oxidation leads to more permanent, irreversible chemical modifications, such as the glyco-oxidation products carboxymethyl lysine and pentosidine, and these products have been used as markers for age-related glycation of proteins. The particular amino-acid side chains involved in AGE reactions depend on the local microenvironment, such as pH and charge. Reactive carbonyl groups for the Maillard reaction may come from molecules other than proteins, such as lipids, nucleic acids and the products of carbohydrate auto-oxidation⁴.

The importance of the Maillard reaction has long been recognized in the food industry, where it causes spoilage and the production of colour and flavour in cooked foods. But its physiological relevance was not evident until after 1977, when Cerami and colleagues identified an Amadori product in a minor haemoglobin variant in diabetes. Since then, glycated proteins such as lens crystallins, collagen and low-density lipoproteins have been implicated in the pathogenesis of late diabetic complications^{4,5}.

AGE-modified proteins are detergent insoluble, protease resistant and have characteristic fluorescent spectra, brown pigmentation and modified amino acids. Several of these features are reminiscent of material isolated from various neuropathological deposits⁶. The stability of plaques and tangles in Alzheimer's disease makes them ideal glycation substrates, and the high lysine content in tau protein (10 per cent of all amino

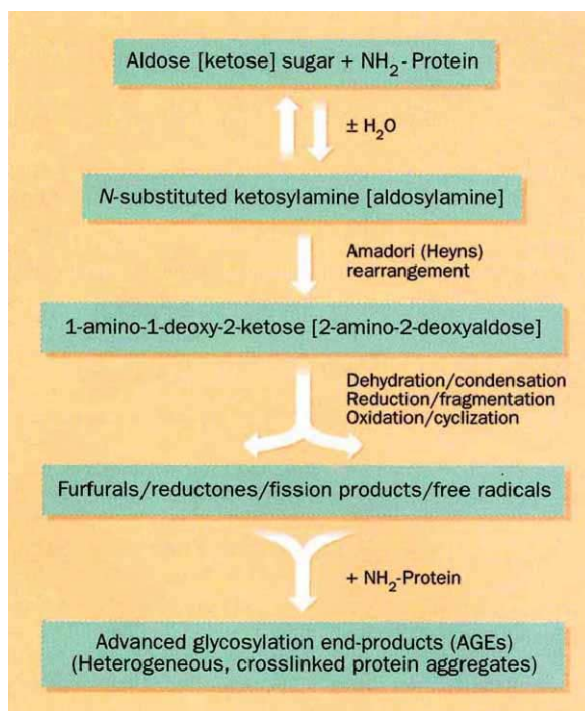
tangles and plaques in Alzheimer brains, and glycated proteins colocalized immunohistochemically with tangles.

Adding to the evidence, two articles in press describe the glycation of tau proteins in tangles. Avila and colleagues⁷ propose that glycation of tau, in addition to hyperphosphorylation, leads to the formation of paired helical filaments in Alzheimer's disease. And Stern and co-workers⁸ have used a further antibody against AGEs to colocalize these AGE-modified proteins with the tau in paired helical filaments. Stern's group postulate that glycation of tau contributes to the stabilization of paired helical filaments within tangles as well as generating reactive oxygen intermediates. Free radicals produced during AGE formation could further promote $A\beta$ fibril aggregation by metal-catalysed oxidation⁹. Inhibition of both protein glycation and oxidant stress, therefore, provides potential sites for preventing the neuronal damage that occurs in Alzheimer's disease. Aminoguanidine has been used already as an inhibitor of AGE crosslinking in experimental diabetes⁴, and amelioration of diabetic nephropathy in mice with an antibody to Amadori products¹⁰ suggests an alternative therapeutic strategy.

What next? Identification of different AGE products and the availability of a range of antibodies against AGEs will help to address the extent of glycation and the nature of the proteins glycated in Alzheimer's disease. Proteins other than tau and $A\beta$ could be glycated and these, in turn, could be involved in further protein crosslinking. It is also not known whether glycation is a secondary event that occurs over a prolonged period to make the proteins gradually more insoluble, or whether it occurs before, and results in, aggregation. In this respect, it should be noted that it is the aggregated form of $A\beta$, rather than the soluble monomer, that is neurotoxic¹.

To what extent do circulating AGEs have a part in Alzheimer's disease pathology? Degradation products of AGE-modified protein are not cleared during renal dialysis of diabetic patients¹¹. These low-molecular-weight AGEs are chemically reactive for further modification of proteins. It will be of interest to find out whether fragmented AGEs are involved in the Alzheimer-type changes in tau protein processing and $A\beta$ deposition that have been observed in the brains of patients undergoing renal dialysis¹².

Finally, there is another connection worth following up. The apolipoprotein E type $\epsilon 4$ allele has been associated with both familial and sporadic late-onset



Protein glycation and AGE product formation.

acids) provides adequate sites for such modification. Furthermore, the normal non-amyloidogenic secretory processing pathway for the $A\beta$ precursor protein cleaves after a lysine residue within the $A\beta$ domain, and modification of this residue would result in abnormal processing¹.

Cerami and co-workers² have found that AGE-modified synthetic $A\beta$ can act as a 'seed' for the accelerated aggregation of soluble $A\beta$ *in vitro*. Moreover, the AGE adducts measured using an antiserum specific for 'late' AGEs were increased threefold in plaque fractions isolated from the cortex of Alzheimer patients compared with those from control brains (the particular AGE species that enhances this nucleation remains to be determined). In the second report³, two Maillard products (pyrraline and pentosidine) not recognized by the antisera used by Cerami *et al.* have been immunohistochemically identified in both

Alzheimer's disease¹³, and apoE type $\epsilon 2$ seems to protect against the disease¹⁴. The association with apolipoprotein polymorphisms and the selective binding of different apoE isoforms to $A\beta$ (ref. 15) are intriguing for several reasons. Clearance of glycated products in the liver is assisted by apolipoproteins and occurs largely in circulating macrophages and the Kupffer cells in the liver, which have specific receptors for AGE-modified proteins^{4,5}. In the brain, apolipoprotein E is produced predominantly in astrocytes and microglia, which are the resident brain macrophages. These macrophages are largely sessile, so it is conceivable that the clearance of AGE-modified proteins is less efficient in the brain. Furthermore, the binding of apoE to $A\beta$ is enhanced under oxidative conditions¹⁵ that can be generated during the formation of AGEs. ApoE itself could be susceptible to glycation. The three main apoE isoforms vary by amino-acid substitutions at two positions: whereas apoE2 has two cysteine residues at positions 112 and 158, these are both substituted with arginine in apoE4. Because arginine can be modified by AGEs, the association of Alzheimer's disease with $\epsilon 4$ and the protection afforded by the $\epsilon 2$ allele could be related in some way to preferential glycation of the apoE4 isoform.

Glycation affects a broad range of macromolecules (proteins, fatty acids, carbohydrates and nucleic acids). Much of our understanding of its pathophysiological relevance has stemmed from investigations into diabetes. That experience should prove invaluable in further investigations of this process in Alzheimer's disease. □

Charles R. Harrington is at the Cambridge Brain Bank Laboratory, MRC Centre, Hills Road, Cambridge CB2 2QH, UK. Camilo A. L. S. Colaco is at the Quadrant Research Foundation, Maris Lane, Trumpington, Cambridge CB2 2SY, UK.

1. Terry, R. D., Katzman, R. & Bick, K. L. (eds) *Alzheimer Disease* (Raven, New York, 1994).
2. Vitek, M. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **91**, 4766–4770 (1994).
3. Smith, M. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **91**, 5710–5714 (1994).
4. Baynes, J. W. & Monnier, V. M. (eds) *The Maillard Reaction in Aging, Diabetes and Nutrition* (Liss, New York, 1989).
5. Vlassara, H., Bucala, R. & Striker, L. *Lab. Invest.* **70**, 138–151 (1994).
6. Colaco, C. A. L. S. & Harrington, C. R. *Neuro Report* **5**, 859–861 (1994).
7. Ledesma, M. D., Bonay, P., Colaco, C. & Avila, J. J. *biol. Chem.* (in the press).
8. Yan, S.-D. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
9. Dyrks, T., Dyrks, E., Hartmann, T., Masters, C. L. & Beyreuther, K. *J. biol. Chem.* **267**, 18210–18217 (1992).
10. Cohen, M. P., Hud, E. & Wu, Y.-Y. *Kidney Int.* **45**, 1673–1679 (1994).
11. Makita, Z. *et al. Lancet* **343**, 1519–1522 (1994).
12. Harrington, C. R. *et al. Lancet* **343**, 993–997 (1994).
13. Saunders, A. M. *et al. Lancet* **342**, 710–711 (1994).
14. Corder, E. H. *et al. Nature Genet.* **7**, 180–183 (1994).
15. Strittmatter, W. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **90**, 8098–8102 (1993).

MACHOs not so tough

Joachim Wambsganss

LAST autumn, two groups of astronomers^{1,2} announced possible discoveries of microlensing events of stars in the Large Magellanic Cloud (LMC). Both groups followed a strategy originally proposed by Paczyński³, namely to monitor the brightness of more than a million stars in the LMC, searching for the signature of massive compact halo objects (or MACHOs). Hopes were raised that this was, at last, evidence of what makes up the long-sought dark matter in the halo of our Galaxy. But on page 275 of this issue⁴, Sahu provides a new interpretation of these observations: he suggests that ordinary stars inside the LMC itself could have produced the observed microlensing events.

It has been known for some time that disk galaxies (our Galaxy⁵ amongst them) are surrounded by a halo of dark matter. Candidates for this dark matter abound, but can be roughly divided into two classes: known or hypothetical elementary particles (among them the WIMPs, or weakly interacting massive particles), and nonluminous astronomical objects, the so-called MACHOs. The latter can be detected indirectly by their gravitational pull on light rays from background stars. If a compact object passes (as viewed from the Earth) in front of a star in the LMC, then its gravitational lens effect causes the observed brightness of the background star to increase in a very characteristic way³.

A microlensing lightcurve has a few more features that help distinguish it from the very large number of variable stars: it is achromatic, symmetrical and unique (it does not recur). The problem, however, is that such events are so rare: even if the halo of the Galaxy were made completely out of MACHOs, only about one in a million LMC stars would be magnified at any given time.

Identification

The characteristic timescale of any detected microlensing event yields only a combination of three unknown parameters: the distance of the lensing object, its mass, and its velocity transverse to the line of sight. That means that by assuming a certain dynamical model of the Galactic halo, which provides a probability distribution for distances and transverse velocities, one can get a rough estimate for the mass of the object. For the events seen last year, the estimated masses lie between a few hundredths and half that of the Sun. The nature of the compact objects cannot be determined from these observations; they could be ordinary low-mass stars, white dwarfs, black holes or

brown dwarfs (objects with too little mass to start nuclear burning in their cores).

What Sahu⁴ does is to calculate the probability for microlensing of a (background) star in the LMC by a (foreground) star in the bar of the LMC. This, he finds, is about an order of magnitude smaller than the rate expected for a halo made of MACHOs. Sahu points out that the observed event rate seems to be well below the expected rate, and that two out of three events are in or close to the bar. So he suggests that the observed microlensing events can be explained by ordinary stars in the LMC. Does this render all the excitement, hope and anxiety that accompanied the possible discovery of dark matter futile?

Galactic bulge

First, it is worth looking at a very similar experiment that searches for microlensing events in stars of the Galactic bulge, a region very close to the centre of the Galaxy. The idea is identical to that of the LMC experiments, except that microlensing events in the direction of the bulge need not be ascribed to any dark objects, but rather to ordinary stars in the disk of the Galaxy^{6,7}.

The first event in this experiment was announced by a third group in this game at about the same time as the LMC events⁸. By now, they have identified as many as eleven microlensing events in the bulge direction (ref. 9, and A. Udalski, personal communication). (The latest of these was the first event caught in action: thanks to on-line data reduction coupled with an early warning system¹⁰, anyone interested could observe the lensed star, produce a lightcurve or take spectra at different phases of the microlensing event.) One of the LMC groups also looks towards the bulge for part of their observing time, and has now reported four events (C. Alcock *et al.*, manuscript in preparation).

There is a problem with the number of events detected here: there are too many of them, more than the stars in the Galactic disk could produce. It turns out that in

1. Alcock, C. *et al. Nature* **365**, 621–623 (1993).
2. Aubourg, E. *et al. Nature* **365**, 623–625 (1993).
3. Paczyński, B. *Astrophys. J.* **304**, 1–5 (1986).
4. Sahu, K. C. *Nature* **370**, 275–276 (1994).
5. Fich, M. & Tremaine, S. A. *Rev. Astr. Astrophys.* **29**, 409–445 (1991).
6. Paczyński, B. *Astrophys. J.* **371**, L63–66 (1994).
7. Griest, K. *et al. Astrophys. J.* **372**, L79–82 (1994).
8. Udalski, A. *et al. Acta astr.* **43**, 289–294 (1993).
9. Udalski, A. *et al. Acta astr.* **44**, 165–189 (1994).
10. Paczyński, B. *IAU Circ.* 5997 (1994).
11. Kiraga, M. & Paczyński, B. *Astrophys. J.* (in the press).
12. Binney, J. *et al. Mon. Not. R. astr. Soc.* **252**, 210–218 (1991).
13. Blitz, L. in *Back to the Galaxy*, AIP Conf. Proc. 278 (eds Holt, S. S. & Verter, F.) 98–106 (New York, 1993).

fact the stars inside the bulge itself, those that are closer to us, can 'lens' those that are further behind¹¹. And even this is not enough: most likely the bulge of our Galaxy is in fact a bar, with its long axis inclined at about 15° to the line of sight^{12,13}, and this elongated geometry may account for the observed high rate of microlensing events (B. Paczyński *et al.*, manuscript in preparation). The similarity to the LMC case is striking.

Now back to Sahu's suggestion⁴ that stars inside the LMC produce the microlensing events. Does it mean MACHOs are dead, and WIMPs alive again? Not at this point. The nice aspect of Sahu's idea, however, is that it is verifiable. He predicts about an order of magnitude fewer microlensing events for LMC lenses than for MACHOs, and emphasizes that for LMC lenses at least some microlensed lightcurves should slightly deviate from the 'point-source approximation' (the relative closeness of the lens means that

the physical size of the lensed star cannot be neglected).

The strongest prediction of Sahu's idea is the following. Whereas microlensing of LMC stars by MACHOs is equally likely for any LMC star (the expected microlensing rate is proportional to the stellar density), Sahu's proposal would give a much higher event rate towards the centre of the LMC.

With three events reported so far^{1,2}, there is no way to decide yet whether the lenses are dark objects in the halo or ordinary stars in the LMC. But more events will be found, and ultimately we should be able to answer the question of whether the halo of the Galaxy is made of MACHOs. We may just have to be patient for a few more years. □

Joachim Wambsganss is in the Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Strasse 1, 85740 Garching bei München, Germany.

IMMUNODEFICIENCY

Zapping T-cell responses

Roger M. Perlmutter

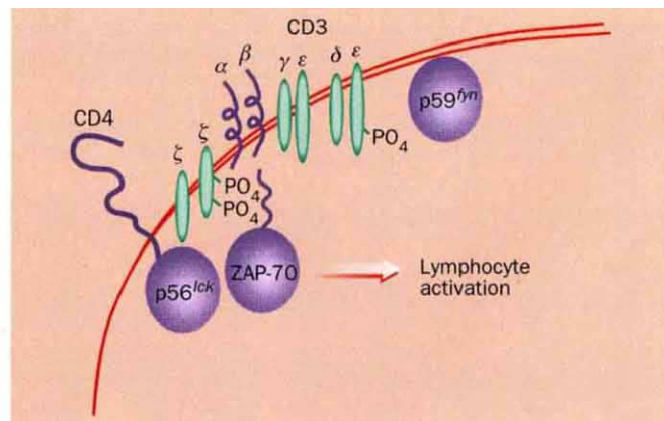
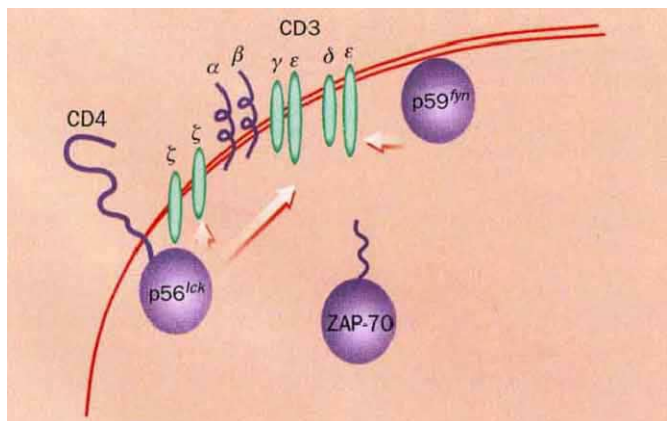
CLINICAL studies of rare human diseases frequently provide unexpected insights into basic biological processes. As an example, three recent papers, one in *Cell*¹ and two others in *Science*^{2,3}, demonstrate that mutations in the aptly named ZAP-70 protein tyrosine kinase gene 'zap' immune function almost completely. Apart from documenting the pivotal role of ZAP-70 in normal T-cell responses, and defining thereby the molecular mechanism underlying one form of autosomal recessive

immunodeficiency disease, the evaluation of children lacking ZAP-70 protein provides a novel perspective on signals that entrain T-lymphocyte development.

ZAP-70 first emerged as a phosphotyrosine-containing protein of relative molecular mass 70K that becomes associated with the ζ -chains of the T-cell antigen-receptor complex (hence the name, zeta-associated protein-70) during T-cell activation⁴. One of a set of structurally related non-receptor protein tyrosine

kinases thought to participate in delivering signals from the T-cell antigen receptor to the cell interior, ZAP-70 attracted special interest when it was shown that aggregation of this kinase, achieved through expression of chimaeric ZAP-70 molecules linked to an extracellular ligand-binding domain, could provoke relatively normal activation responses in T cells⁵. This observation, coupled with previous studies documenting the participation of the *src*-family kinases p59^{lyn} and p56^{lck} in T-cell signalling⁶, encouraged the promulgation of a model in which stimulation of the T-cell antigen receptor initially triggers (presumably via p56^{lck}) tyrosine phosphorylation of specialized motifs in antigen-receptor components, notably the CD3 γ , CD3 δ , and CD3 ϵ chains and the ζ -homodimer. Thereafter, the phosphotyrosine residues on these chains could serve as interaction sites for the specialized phosphotyrosine-binding (SH2) domains of the ZAP-70 kinase which, once aggregated, would elicit subsequent biochemical responses⁷.

This model clearly predicts that T lymphocytes lacking functional ZAP-70 protein should not respond satisfactorily to antigenic stimulation. Arpaia *et al.*¹, seeking to elucidate the biochemical basis of a rare, selective T-cell immunodeficiency, were the first to report results consistent with this view. They observed that circulating lymphocytes in these patients, though normal in number, failed to proliferate following T-cell receptor stimulation. Subsequent studies showed that although p56^{lck} and p59^{lyn} were present in normal amounts in lymphocytes from affected individuals, ZAP-70 protein was undetectable. Elder *et al.*² and Chan *et al.*³ have since identified additional cases of this immunodeficiency, and have together



Proposed involvement of the ZAP-70 protein tyrosine kinase in T-cell activation. The left hand panel shows the T-cell receptor and its associated protein tyrosine kinases at the time of encounter with an appropriate antigen. The receptor itself consists of an $\alpha\beta$ heterodimer (directly involved in recognizing foreign peptide antigens) associated with the γ -, δ - and ϵ -chains of the CD3 complex, and with a pair of disulphide-linked ζ -chains. All three CD3 polypeptides, as well as the ζ -chains, contain 'activation motifs' consisting of paired tyrosine and leucine residues in a precise topological orientation. Antigen recognition provokes phosphorylation of tyrosine residues in these

motifs via a process that may involve the *src*-family kinases p56^{lck} (typically brought to the complex by virtue of its association with the CD4 and CD8 co-receptors) and p59^{lyn} (which interacts directly with components of the antigen-receptor complex itself). In the second step of the activation process (right hand panel), paired phosphotyrosine-binding domains in the ZAP-70 protein tyrosine kinase allow it to interact with the phosphorylated ζ -homodimer. This association stimulates, either directly or indirectly, ZAP-70 kinase activity, which can then act on a variety of targets to provoke the lymphocyte-activation response.

defined three different null alleles of the ZAP-70 gene, each of which apparently directs the synthesis of an unstable protein.

The interpretation of these studies is in principle straightforward: absence of ZAP-70 protein results in an inability to couple the T-cell antigen-receptor complex to downstream signalling pathways, hence only inert T lymphocytes are produced. Normal proliferation occurs in T cells lacking ZAP-70 when stimulated with agents (ionomycin and phorbol esters) known to bypass the antigen receptor; this provides some support for the view that T cells from such individuals fail only in the earliest stages of antigen-receptor-induced signal transduction. But the most remarkable feature of children who inherit ZAP-70 mutations is a peculiar skewing in T-cell production, such that only CD4⁺, and not CD8⁺, cells emerge from the thymus to populate peripheral lymphoid organs. So there is reason to wonder whether the lymphocytes that circulate in such patients are products of a normal maturation pathway.

T lymphocytes ordinarily mature and leave the thymus only after receiving an antigen-receptor-initiated stimulus, a phenomenon referred to as positive selection⁸. In patients deficient in ZAP-70, analysis of thymic biopsies suggests that T-cell maturation arrests precisely at the point where positive selection is required¹. This would make sense if ZAP-70 acts by transmitting the antigen receptor signal required for selection to proceed. How then to explain the presence of circulating CD4⁺ T cells in these patients?

At present, there is no basis for discriminating between the intracellular signalling pathways required for development of CD4⁺ as opposed to CD8⁺ T cells, although there is at least one mutation in a mouse transcription factor gene that affects these populations differently⁹. If the CD4⁺ lymphocytes in patients deficient in ZAP-70 traverse a normal differentiation pathway to gain access to the circulation, then signalling from the antigen receptor must proceed to some extent in the complete absence of ZAP-70 protein. Alternatively, if the CD4⁺ cells in these patients arrive in the circulation by an aberrant maturation pathway, then the defect in T-cell receptor signalling observed in such cells may not result

solely from the absence of ZAP-70.

Although detailed biochemical studies and the development of appropriate mouse model systems will be required before ZAP-70 can be placed in its appropriate context, the recent identification of patients with ZAP-70 mutations further attests to the value of clinical investigation in basic biological discovery. In addition, although ZAP-70 defects may prove an uncommon cause of hereditary immunodeficiency, treatment of these patients by gene replacement may be especially effective, as reconstitution of ZAP-70 gene expression should yield cells with an inherent selective advantage.

Finally, examination of patients deficient in ZAP-70 has rekindled interest in non-receptor protein tyrosine kinases as targets for pharmacological intervention. Because defects in ZAP-70 synthesis selectively compromise T-cell function, agents designed to block ZAP-70-mediated catalysis should effectively terminate autoimmunity, and could perhaps achieve this result without the toxic side effects that complicate currently available immunosuppressive regimens. □

Roger M. Perlmutter is in the Department of Immunology, University of Washington, Seattle, Washington 98195, USA.

PSYCHOLOGY

Comprehending baby-think

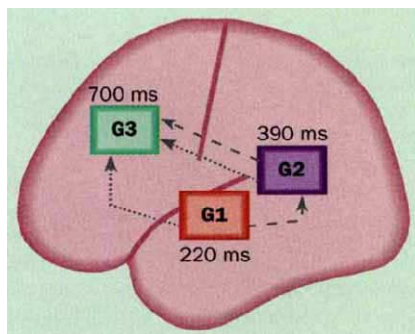
Anne Christophe and John Morton

How can we capture the rich cognitive life of young infants? Behavioural observations are restricted by the limited repertoire of infant responses: they can look and suck, and that is about it. Although

temporally and spatially separate responses to be identified that correspond to three separate processes. Moreover, the infant brain takes less than 400 milliseconds to discriminate a phonetic contrast.

Phonetic perception over the first year of life has been particularly well studied behaviourally: surprisingly, infants during the first few months of life can discriminate all the phonetic contrasts, irrespective of whether they are actually used in their native language. Thus, infants must start with a universal representation, that is, one appropriate for all the languages in the world. They end up as adults with a language-specific representation. For instance, Japanese adults, unlike Japanese infants, cannot hear the difference between 'r' and 'l', two sounds that don't have a separate existence in Japanese. We know nothing of the structural changes underlying such behavioural development. It is possible that the universal format is destroyed as the infant consolidates the contrasts of the native language. Alternatively, the universal format may be preserved for a while but be inhibited by the language-specific one. These questions can hardly be addressed with standard behavioural discrimination techniques.

The high-density ERP method³ has been successfully adapted to two-month-old infants by Dehaene-Lambertz and Dehaene. Compared with other brain-imaging techniques, it has the advantage of being non-invasive: a light net of wet electrodes is simply laid on the infant's head (see front cover). Compare this to a PET scan, in which radioactivity has to be injected. The ERP method relies on the fact that a burst of localized activity in the cortex acts as a current 'generator', producing an electrical field that can be picked up by electrodes on the scalp as



Suggested sequence of information flow in the brain. The boxes G1 to G3 represent the approximate location of the inferred generators underlying the three peaks of electrical activity picked up on the surface of the skull, together with the average time of occurrence of the peaks from stimulus onset. We assume that each generator corresponds to a neural network performing a particular computational function. Dashed and dotted arrows denote two possible hypotheses concerning the information flow between the generators.

these responses have told us a great deal about the cognitive capacities of infants¹, to learn more we have to investigate more directly. On page 292 of this issue² Dehaene-Lambertz and Dehaene use the high-density, event-related potential (ERP) technique adapted to two-month-old babies to find out how they process simple syllables. One undifferentiated response is as much as could be hoped for from behavioural measures with infants of this age. The high-density ERP technique, combined with a simple and ingenious experimental design, has enabled three

1. Arpaia, E., Shahar, M., Dadi, H., Cohen, A. & Roifman, C. M. *Cell* **76**, 947–958 (1994).
2. Elder, M. E. *et al.* *Science* **264**, 1596–1599 (1994).
3. Chan, A. C. *et al.* *Science* **264**, 1599–1601 (1994).
4. Chan, A. C., Iwashima, M., Turk, C. W. & Weiss, A. *Cell* **71**, 649–662 (1992).
5. Kolanus, W., Romeo, C. & Seed, B. *Cell* **74**, 171–183 (1993).
6. Perlmutter, R. M., Levin, S. D., Appleby, M. W., Anderson, S. J. & Alberola-Lia, J. A. *Rev. Immun.* **11**, 451–499 (1993).
7. Weiss, A. & Littman, D. R. *Cell* **76**, 263–274 (1994).
8. Marrack, P. & Parker, D. C. *Nature* **368**, 397–398 (1994).
9. Matsuyama, T. *et al.* *Cell* **75**, 83–96 (1993).

changes in potential. Temporal resolution of this kind of surface recording is not new; what is new is that, because of the high density of the electrodes, the location of the generators can be inferred much more precisely from the pattern of surface activity — and that this can readily be accomplished with infants.

The new application of this technique looks promising. The authors presented infants with a series of four identical syllables (the standard), followed by a fifth that was either identical or phonetically different (deviant). They observed three distinct responses, or peaks of electrical activity, to these syllables. Their Fig. 3 shows the pattern of electrical activity on the surface of the head for the first two peaks: the anterior positivity synchronous with posterior negativity reflects the presence of generators within the temporal lobes. Apart from being distinct temporally and spatially, these two generators differ functionally. This can be seen in the responses to the fifth syllable, where peak 2, unlike peak 1, discriminates between phonetically different syllables. The final peak, seen as an anterior negativity, occurs only when the deviant syllable is presented. Because the negativity was rather diffuse over the centro-frontal regions, its generator could not be accurately localized. But the authors make a good argument that it may arise from an anterior cortical or subcortical neural circuit (such as the anterior cingulate cortex).

What can we say about the three sources of activity? One interpretation, perhaps the most intuitive, is that information flows from one brain area to the next in a sequence corresponding to the temporal sequence (see figure, dashed arrows), but this is not the only interpretation. For example, there could be a direct connection between G1 (the generator underlying Dehaene-Lambert's and Dehaene's peak 1) and G3 (dotted arrows). To test this, we would look for a manipulation that would cause G1 but not G2 to dishabituate, and see whether G3 followed it. The most obvious option is that G1 is sensitive to changes in intensity or pitch whereas G2, being dedicated to phonetic processing, is insensitive to such changes. In this case, if G3 were connected to both G2 and G1 then it would respond to prosodic changes as well as phonetic ones.

It would be exciting to track the evolution of these responses over the first year of life. Thus, presenting Japanese babies with a 'ra-la' contrast, we expect to see a loss of dishabituation in at least one

cortical response at the end of the first year, corresponding to the disappearance of the behavioural response to this contrast. The interesting thing would be if some other cortical response still dishabituates to this, non-native, contrast, indicating that some universal processing

is still preserved. The high-density ERP technique would readily lend itself to answering such questions. □

Anne Christophe and John Morton are at the MRC Cognitive Development Unit, 4 Taverton Street, London WC1H 0BT, UK.

INFLAMMATION

Key mediator takes shape

Nancy A. Thornberry

No one can dispute the importance of interleukin-1 β converting enzyme (ICE). Formerly known only as the enzyme responsible for the production of interleukin-1 β , a key mediator of inflammation, it has recently also been implicated in programmed cell death. The crystal structure of this protease, described by Wilson *et al.* on page 270 of this issue¹, puts the relationship of ICE to the products of known cell-death genes on a secure footing, and provides a molecular framework for the rational design of inhibitors that may prove crucial in unravelling the enzyme's biological functions.

Interleukin-1 (IL-1) is the general term for two monocytic proteins, IL-1 α and IL-1 β , which bind to the same receptor with comparable affinity. It has long been considered an attractive target for therapeutic intervention in chronic and acute inflammation. The discovery of a naturally occurring IL-1 receptor antagonist, IL-1ra, which is effective in several animal models of inflammation, has justified this interest². But attempts to find a low-molecular-weight receptor antagonist have proved unfruitful, prompting the search for alternative strategies. Meanwhile, investigators of IL-1 β found that it is synthesized as an inactive precursor of relative molecular mass 31K which requires proteolytic cleavage at an Asp-Ala site to produce the 17.5K mature, biologically active form. On its discovery six years ago, the enzyme responsible for this unusual cleavage (ICE) immediately presented itself as an attractive therapeutic target^{3,4}.

Allure

This enzyme became even more alluring following its purification and cloning in 1992, which indicated that it was not related to any known protease^{5,6}. The active enzyme was shown to require two subunits (of 10K and 20K), both of which are autoproteolytically derived from an inactive 45K proenzyme. Mechanistic studies identified the enzyme as a cysteine protease, led to the design of a potent peptide aldehyde inhibitor ($K_i = 0.76$ nM), and indicated that the catalytic thiol was at position 285. All of this is now confirmed by the crystal structure, which

shows the same tetrapeptide aldehyde covalently bound to Cys 285. As with the catalytic groups found in other cysteine proteases, a histidine (His 237) seems to serve as a general base, and the proposed overall mechanism is in accord with current thinking in the protease field. The structure also shows that both subunits of the enzyme contribute residues to the active site, explaining why the monomers are individually inactive.

Pointers

The authors have kept several details of the interaction between the inhibitor and the active site to themselves, but the crystal structure nonetheless provides helpful pointers for those contemplating design of non-peptide inhibitors. It defines the important interactions of the inhibitor side chains to be with the S₁ and S₄ subsites, as expected from the previously reported substrate specificity of the enzyme. The structure also shows the tetrapeptide aldehyde, a potential transition state analogue, bound in a non-transition-state conformation, with the oxyanion of the inhibitor being stabilized by the active site His 237. This observation is not only of theoretical interest to those trying to reconcile conflicting data about the way in which peptide aldehydes bind in other cysteine proteases, it is also of practical importance: ICE may be more tolerant of non-peptide structural modifications to an inhibitor that is not a transition-state analogue. Finally, in the crystal two heterodimers associate to form a tetramer, and it has been argued that this is the catalytically active form of the enzyme. If this is true, the interface between these heterodimers is another possible target for inhibitor design.

Beyond the development of a selective, non-peptide inhibitor, a number of serious questions regarding the biology of ICE remain. For example how is IL-1 β released from stimulated monocytes? Electron microscopy indicates that the active enzyme is localized on the membrane⁷, but the authors make no mention of a possible membrane-binding domain. If the active enzyme is indeed membrane-bound, then it may be involved in both processing and release of

1. Mehler, J. & Dupoux, E. *What Infants Know* (Blackwell, Oxford, 1994).
2. Dehaene-Lambertz, G. & Dehaene, S. *Nature* **370**, 292–295 (1994).
3. Tucker, D. *Electroenceph. clin. Neurophysiol.* **87**, 154–163 (1993).

IL-1 β . Another concern is that the vast majority (99 per cent or more) of enzyme in both stimulated and resting monocytes is in the form of the inactive proenzyme, and no active enzyme can be detected by standard biochemical methods⁸. How is enzyme regulated in these cells? Most importantly, will inhibition of IL-1 β alone be effective in treating chronic or acute inflammation (or both), or is IL-1 α also a central player?

As if IL-1 biology were not challenging enough, late last year ICE was shown to have sequence similarity to CED-3 protein, the product of a gene required for programmed cell death in *Caenorhabditis elegans*⁹. From the crystal structure we see that all of the residues implicated in catalysis and binding, including the three amino acids important for recognition of the P₁ aspartic acid, are conserved between ICE and CED-3. This leaves little doubt that *ced-3* encodes a cysteine protease which cleaves after an aspartic acid in its (as yet unidentified) substrate or substrates.

That ICE itself is involved in programmed cell death is not as certain. The most compelling biological evidence implicating the enzyme has been provided by a potent viral serpin inhibitor, *crmA* protein, which prevents cell death when expressed in neurons of the dorsal root ganglion¹⁰. This is inconclusive, because cell death could be due to an ICE homologue that is also inhibited by *crmA* protein, like the broad inhibition of serine proteases by α -1-proteinase inhibitor¹¹. The existence of at least one human homologue is virtually assured, given the identification of the murine *nedd2* gene, which shows conservation of all the binding and catalytic residues common to ICE and CED-3 (ref. 12). Potent and selective inhibitors rationally designed with the X-ray structure should lead to a greater understanding of ICE and its relatives in inflammation and programmed cell death. □

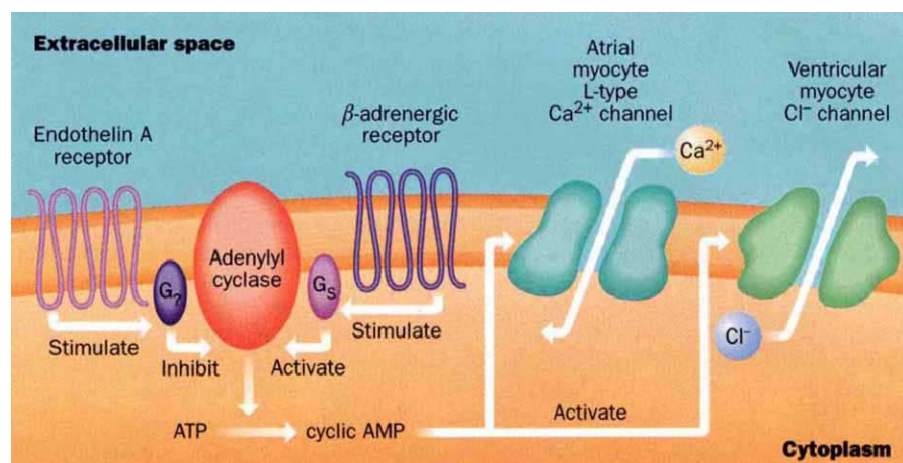
Nancy A. Thornberry is in the Department of Enzymology, Merck Research Laboratories, 126 East Lincoln Avenue, Rahway, New Jersey 07090, USA.

Endothelin to the rescue?

Edgar Haber and Mu-En Lee

THE peptide hormone endothelin powerfully increases the contractility of both smooth and cardiac muscle. Now two reports^{1,2} on pages 297 and 301 of this issue on the effects of endothelin on ion channels give new insight into how it may be able to protect heart muscle undergoing ischaemia against the shortening of the action potential by catecholamines¹. The hormone is shown to inhibit the protein kinase A-dependent chloride current in ventricular myocytes¹ and also to

domain receptors^{7,8}, ET_A and ET_B. ET_A is selective for endothelins, binding ET-1 with slightly greater affinity than ET-2, and ET-2 much more strongly than ET-3, but ET_B is promiscuous. The initial idea that endothelins are secreted solely by endothelial cells and effect contraction only in vascular smooth muscle has been broadened to take account of the fact that they can cause the release of nitric oxide and consequent vasodilatation. They also increase cardiac contractility and, like



Counterbalancing effects of the endothelin A (ET_A) receptor and the β -adrenergic receptor on cyclic AMP production and consequent flux of Ca²⁺ and Cl⁻. G_i represents a pertussis-toxin-sensitive GTP-binding protein.

hyperpolarize the membrane in atrial myocytes by inhibiting the L-type Ca²⁺ channel and activating the K⁺ channel². By opposing the adverse effect of catecholamines on cardiac electrophysiology, endothelin may moderate the risk of dangerous ventricular arrhythmias which often occur in the course of a heart attack.

A curious abstract³ was noticed by members of Tomoh Masaki's laboratory (Masashi Yanagisawa, personal communication). This abstract described how the supernatant from cultured endothelial cells caused rings cut from arteries to contract; this effect, attributed to a molecule of relative molecular mass 85K, was destroyed by proteolytic enzymes, suggesting that an unknown vasoactive peptide was in operation. The Masaki group was prompted to isolate and characterize this effector and soon found three structurally and pharmacologically distinct isopeptides of 21 amino-acid residues apiece — the endothelins ET-1, 2 and 3 — which are each encoded by different genes^{4,5}.

The mature, active endothelins are cleaved from a precursor of about 40 residues by a specific metalloprotease converting enzyme⁶. The peptides act on two distinct, seven-transmembrane-

many cytokines, can act on an array of tissues (lung, kidney, central nervous system) and promote mitogenesis. Indeed, homozygous disruption of the ET-1 gene is incompatible with life⁹, not because of the expected malfunction of the circulation but because of respiratory failure.

Intracellular signalling by the endothelin receptors is complex. While the ET_A and ET_B receptors both stimulate phosphatidylinositol hydrolysis and arachidonic acid release, the effects of the two receptors on intracellular cyclic AMP production vary according to the cell type studied¹⁰. The endothelins increase inwardly rectifying K⁺ currents in atrial cells¹¹, activate low conductance Cl⁻ channels in aortic smooth muscle cells¹², and activate the voltage-dependent Ca²⁺ channel in vascular smooth muscle cells¹³. Now James *et al.*¹ extend these and related observations to ventricular myocytes by showing that ET-1 (acting through the ET_A receptor) inhibits adenyl cyclase through a pertussis-toxin-sensitive GTP-binding protein, thereby inhibiting the Cl⁻ channel and lengthening the action potential (see figure). This work also has clinical implications, for myocardial infarction is associated with often fatal cardiac arrhythmias that are mediated by catechol-

1. Wilson, K. P. *et al.* *Nature* **370**, 270–275 (1994).
2. Dinarello, C. A. & Thompson, R. C. *Immun. Today* **12**, 404–410 (1991).
3. Black, R. A., Kronheim, S. R. & Sleath, P. R. *FEBS Lett.* **247**, 386–390 (1989).
4. Kostura, M. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **86**, 5227–5231 (1989).
5. Cerretti, D. P. *et al.* *Science* **256**, 97–100 (1992).
6. Thornberry, N. A. *et al.* *Nature* **356**, 768–774 (1992).
7. Singer, I. I. *et al.* *Lymphokine Cytokine Res.* **12**, 340 (1993).
8. Ayala, J. M. *et al.* *J. Immun.* (in the press).
9. Yuan, J., Shaham, S., Ledoux, S., Ellis, H. M. & Horvitz, H. R. *Cell* **75**, 641–652 (1993).
10. Miura, M., Zhu, H., Rotello, R., Hartwig, E. A. & Yuan, J. *Cell* **75**, 653–660 (1993).
11. Beatty, K., Beith, J. & Travis, J. J. *biol. Chem.* **255**, 3931–3934 (1980).
12. Kumar, S., Kinoshita, M., Noda, M., Copeland, N. G. & Jenkins, N. A. *Genes Dev.* (in the press).

amines, whose concentrations increase during ischaemia. In fact, treatment with β -adrenergic-receptor blockers, which inhibit the action of catecholamines, is now routinely used after heart attacks. As the concentration of ET-1 also increases during myocardial ischaemia, ET-1 may directly counter the electrophysiological effects of catecholamines by preventing the arrhythmogenic shortening of the action potential.

A note of caution must be expressed, however, about extrapolating selective studies in guinea-pig and rabbit myocytes to humans. The studies by James *et al.*¹ and Ono *et al.*² focus largely on ET-1 acting through the ET_A receptor. Molenaar *et al.*¹⁴ have shown that ET_B receptors are more prevalent than ET_A recep-

tors in the atrioventricular node and the bundle of His, two potential sites of origin for ventricular arrhythmias. Stimulation of ET_B receptors is likely to produce electrophysiological effects very different from those described for ET_A receptors. Moreover, there is great variability among species in the effects of ET-1 on a number of electrophysiological parameters¹⁵. Establishment of the clinical relevance of these new observations must therefore await the results of pharmacological studies on humans. □

Edgar Haber and Mu-En Lee are at the Center for the Prevention of Cardiovascular Disease, Harvard School of Public Health, 677 Huntington Avenue, Boston, Massachusetts 02115, USA.

documented by Seewald, nor for his observation of the frenetic trading of oxygen and hydrogen among the different players in his chemical system as they scurry to maintain thermodynamic equilibrium in response to changing chemical conditions.

Yet Seewald's conclusions are based on hard experimental data; furthermore, they have a theoretical foundation in the work of Helgeson and co-workers⁴, and of Sato⁵, both of which give thermodynamic treatments of water, hydrocarbons and the minerals found in the rocks of commercial hydrocarbon fields. Helgeson and co-workers⁴ linked metastable thermodynamic equilibrium between hydrocarbons, water, oxidized aqueous carbon species such as CO₂ and carboxylic acids, and some mineral species to an overall set of chemical reactions including redox reactions and the 'hydrolytic disproportionation' of petroleum.

Seewald's observations are expected by these treatments, in which the principal control of organic metamorphism is decreasing oxygen fugacity (or because of the ubiquitous presence of water, increasing hydrogen fugacity). With increasing burial, ever lower oxygen fugacities drive organic metamorphic reactions towards completion. Increasing burial temperatures are thought only to promote the thermodynamically favoured reactions and attainment of equilibrium.

Seewald's data are but one more brick of experimental evidence in the building of an alternative model of organic metamorphism. However, previously much of this experimental evidence has gone unappreciated, sometimes even by the original investigators themselves. Hesp and Rigby⁶ demonstrated in 1973 that the presence of water dramatically retarded the thermal destruction of an Australian crude oil. French⁷, during experiments on the synthesis and stability of the mineral siderite (iron carbonate), inadvertently produced measurable amounts of oxidized hydrocarbons (organic acids, alcohols and ketones; and no doubt petroleum hydrocarbons, although French did not report them). These reports of the generation of 'oxidized' hydrocarbons (and probably true hydrocarbons) by rock-water reactions, some more than 30 years ago, may put the 'revolutionary' aspect of Seewald's data in better perspective. And more recently, other groups^{8,9} have inadvertently generated petroleum-like hydrocarbons from water and acetic acid whilst examining thermal stability of the acetic acid in aqueous systems.

Perhaps the most telling experiments, had we realized it, were those of Hoering. In experiments with deuterated water and kerogen¹⁰ and oil shales¹¹, he demonstrated that kerogen (the solid polymeric matter in shales which generates hydrocarbons) can exchange hydrogen with

1. James, A. F., Xie, L.-H., Fujitani, Y., Hayashi, S. & Horie, M. *Nature* **370**, 297–300 (1994).
2. Ono, K. *et al.* *Nature* **370**, 301–304 (1994).
3. Agricola, K. M., Rubanyi, G., Paul, R. J. & Highsmith, R. F. *Fed. Proc.* **43**, 899 (1984).
4. Yanagisawa, M. *et al.* *Nature* **332**, 411–415 (1988).
5. Inoue, A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **86**, 2863–2867 (1989).
6. Oppenorth, T. J., Wu-Wong, J. R. & Shiosaki, K. *FASEB J.* **6**, 2653–2659 (1992).
7. Arai, H., Hori, S., Aramori, I., Ohkubo, H. & Nakanishi, S. *Nature* **348**, 730–732 (1990).

8. Sakurai, T. *et al.* *Nature* **348**, 732–735 (1990).
9. Kurihara, Y. *et al.* *Nature* **368**, 703–710 (1994).
10. Aramori, I. & Nakanishi, S. *J. biol. Chem.* **267**, 12468–12474 (1992).
11. Kim, D. *Circ. Res.* **69**, 250–255 (1991).
12. Van Renterghem, C. & Lazdunski, M. *Pflügers Arch.* **425**, 156–163 (1993).
13. Goto, K. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **86**, 3915–3918 (1989).
14. Molenaar, P. *et al.* *Circ. Res.* **72**, 526–538 (1993).
15. Kohmoto, O. *et al.* *Am. J. Physiol.* **265**, H793–800 (1993).

PETROLEUM GEOCHEMISTRY

Metamorphic free-for-all

Leigh C. Price

ON page 285 of this issue, J. S. Seewald¹ presents results of hydrothermal experiments on systems containing rock, hydrocarbons and water which demonstrate attainment of metastable thermodynamic equilibrium between the three separate phases. His data may provide the missing piece of a changing picture of organic metamorphism and thus oil and gas formation.

Seewald's experiments were done in flexible gold bags containing water and a naturally occurring mineral assemblage which buffered the oxygen fugacity, and therefore the redox potential, of the system. Hydrogen fugacity was also buffered because water disproportionates to oxygen and hydrogen. When experimental conditions were changed, by injecting either ethane or ethene into the system or by changing the temperature, the ethane/ethene ratio in the system gradually readjusted towards the values expected from thermodynamic calculations for the experimental conditions. Seewald correctly concludes that ethane, ethene, water and the mineral assemblage buffering the redox potential of the system are all in reversible equilibrium, and exchange hydrogen (or oxygen) with one another.

To physical scientists in general, Seewald's data may seem expected, even ordinary. But the data are very exciting to

some of us working in the obfuscated fields of petroleum geology and geochemistry where system controls in nature remain obscure and where laws applicable to the other physical sciences appear occasionally and inexplicably suspended in accepted models. To appreciate this point, one must understand conventional wisdom concerning, first, the degree with which rocks and water interact with hydrocarbons, and second, the process controls in models for organic metamorphism.

By accepted thought, aside from the precipitation of carbonate-bearing minerals caused by the copious generation of carbon dioxide along with hydrocarbons in source rocks, mineral species and water simply do not interact with hydrocarbons, *finio*. That hydrogen and oxygen would be traded like baseball cards among water, hydrocarbons and mineral species is a concept which some investigators in this field will find preposterous.

Oil and gas (hydrocarbon) generation is conventionally held to occur by irreversible first-order reactions where geological time and burial temperature are the only controls and may be substituted for each other in Arrhenius equations describing these reactions^{2,3}. There is no latitude for organic metamorphic reactions displaying the reversibility

water¹⁰. He also demonstrated (by deuterating *n*-docosane, the *n*-C₂₂ *n*-paraffin) that water and pure hydrocarbons exchange hydrogen¹¹. Thus, as Seewald suggests, the hydrocarbon-generating capacity of source rocks may increase as their kerogen incorporates hydrogen from water.

Our own data from steam pyrolysis experiments¹² on organic-rich shales bear this out: 4–8 times as much CO₂ is produced during hydrocarbon generation than could possibly arise from the original oxygen content of the kerogen. The only source for this excess oxygen is water. Our measurements demonstrate that the hydrogen content of the kerogen also increases by this process. Lewan¹³ also noted excess CO₂ production in his hydrous pyrolysis (water-bearing) experiments on oil source rocks and attributed this to oxygen from water incorporated into the kerogen.

Like Seewald, L. M. Wenger and I have seen characteristics of reversibility of organic-metamorphic reactions in our experiments. For example, increasing fluid pressure retards all aspects of organic metamorphism¹⁴. By the theory of Sato and of Helgeson and co-workers, product escape should foster reaction extent, whereas product retention should impede it^{4,5}. Wenger and I have also observed this. Clearly there is considerable support for Seewald's data and conclusions.

Some investigators will dismiss Seewald's study as inappropriate because of the high experimental temperatures. Experiments at lower temperatures for longer times are certainly in order, but Seewald's temperatures are fairly typical of experiments simulating hydrocarbon generation in the laboratory. Furthermore, the participation of ethene in Seewald's reactions is most impressive because carbon-carbon double bond strengths are much higher than those of carbon-carbon single bonds, 145.8 kcal mol⁻¹ rather than 82.6 kcal mol⁻¹.

The origin of the high concentrations of carboxylic acids in petroleum reservoir waters warrants comment. Most of these acids are likely to be byproducts from thermophilic sulphate-reducing bacteria slowly consuming hydrocarbons at the oil-water interface in oil reservoirs over geological time. This fact, not generally appreciated, is not academic, because these microbes have an agenda of their own apart from thermodynamic equilibrium.

The probability of hydrolytic disproportionation carries profound implications. Some are economic, and three of these stand out. First, the principal purpose of petroleum geochemical models is to aid in the discovery and commercial recovery of oil and gas. Existing models of petroleum formation do not account for either hydrolytic disproportionation or hydrogen

enrichment from water and thus may fail to describe the realities of nature. Second, as Seewald points out, the initial hydrogen-to-carbon ratio of kerogen will not be an accurate measure of its ultimate hydrocarbon generation capacity if substantial amounts of hydrogen from water are incorporated into kerogen during progressive burial. Third, in hot, deep-basin conditions the hydrogen-to-carbon ratios of 'spent' kerogen are low, but the kerogen may still have significant gas-generation capacity from incorporation of water, promoted by the high burial temperatures. So perhaps we should look upon 'spent' deep-basin kerogen as a possible gas-generating machine which has yielded much larger in-place gas resources than previously recognized.

Organic metamorphism has tended to be mathematically modelled by analyses from open-system, dry pyrolyses. This approach is particularly ill-suited to modelling natural hydrocarbon generation, which occurs in pressurized water-wet systems. Furthermore, in nature, oxygen fugacities are low and product escape is difficult because the system is closed (or semi-closed) to fluid flow, yet in these analyses product escape occurs instantly and oxygen fugacities are high. I believe a more inappropriate instrument with which to model petroleum generation could hardly be designed.

In my opinion, shared with a limited number of other researchers, petroleum formation has been trivialized by simplistic models which often simulate natural processes poorly¹⁵. One can hope, though, that other well controlled benchmark experiments like Seewald's, and those which have been suggested by Helgeson and co-workers⁴, will lead to realistic models. □

Leigh C. Price is in the Branch of Petroleum Geology, US Geological Survey, MS-940, Denver Federal Center, Denver, Colorado 80225, USA.

1. Seewald, J. S. *Nature* **370**, 285–287 (1994).
2. Hunt, J. M. *Petroleum Geochemistry and Geology* (Freeman, San Francisco, 1979).
3. Tissot, B. P. & Welte, D. H. *Petroleum Formation and Occurrence* (Springer, New York, 1984).
4. Helgeson, H. C., Knox, A. M., Owens, C. E. & Shock, E. L. *Geochim. cosmochim. Acta* **57**, 3295–3339 (1993).
5. Sato, M. in *Fluid–Mineral Interactions: A Tribute to H. P. Eugster* (eds Spencer, R. J. & Chou, I.-M.) 271–283 (The Geochemical Society, San Antonio, 1990).
6. Hesp, W. & Rigby, D. *Erdöl Kohle-Erdgas* **26**, 70 (1973).
7. French, B. thesis, Johns Hopkins Univ. (1964).
8. Palmer, D. A. & Drummond, S. E. *Geochim. cosmochim. Acta* **50**, 813–823 (1986).
9. Bell, J. L. S., Palmer, D. A., Barnes, H. L. & Drummond, S. E. *Geochim. cosmochim. Acta* (in the press).
10. Hoering, T. *Carnegie Inst. Wash. Yb.* **67**, 199 (1968).
11. Hoering, T. *Org. Geochem.* **5**, 267–278 (1984).
12. Wenger, L. M. & Price, L. C. in *Organic Geochemistry—Advances and Applications in Energy and the Natural Environment* (ed. Manning, D. A. C.) 335–339 (1991).
13. Lewan, M. D. *Am. Chem. Soc. Preprints, Division of Fuel Chemistry* **37**, 1643–1649 (1992).
14. Price, L. C. & Wenger, L. M. *Org. Geochem.* **19**, 141–159 (1992).
15. Price, L. C. *Geochim. cosmochim. Acta* **57**, 3261–3280 (1993).

Travelling light

MODERN optical fibre has a lower refractive index at the surface than in the centre. This keeps the beam energy away from the surface so that it cannot leak out. Daedalus is now extending this principle. He points out that if the index within the fibre declines inversely as the hyperbolic cosine of the radius, the fibre acts as a lens. Light falling on its end face is focused to a point inside it. The eyes of certain fishes seem to work this way.

Imagine, says Daedalus, an extended length of such a fibre. Parallel light falling on its end face will form an inverted image one focal length inside the fibre. But its optical odyssey will not end at that point. It will continue to travel down the fibre, being refracted by the same law as it goes. Diverging from its first focus, it will revert to parallel light two focal lengths down the fibre. At three focal lengths it will form an upright real image, and at four focal lengths it will recover the exact direction it had when it went in. The cycle will repeat indefinitely.

Thus this cunning fibre is more than a mere light-guide. It conveys not only light intensity, but image information as well. Cut to an exact multiple of four focal lengths, a bundle of such fibres will accept a pattern of light rays at one end, and emit it undeviated at the other. No matter how long and winding the bundle, it will behave like a thin, transparent window. Look into it, and you will see the scene at the other end.

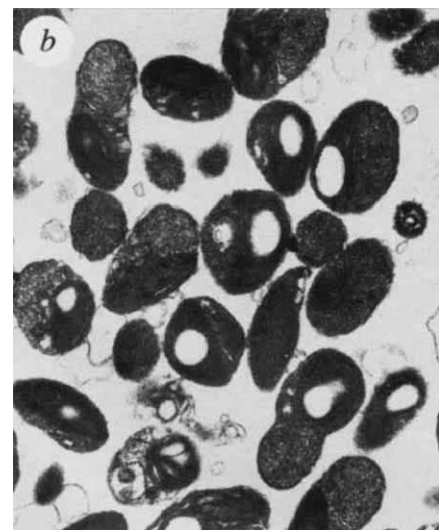
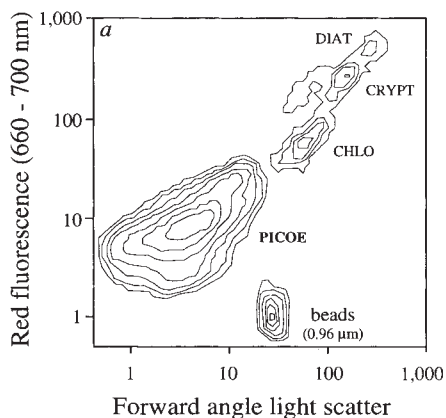
Daedalus is now refining the standard process for making radially index-graded fibre so as to produce fibre of the exact inverse 'cosh law' required. When perfected, 'Coshfibre' will have many uses, especially in remote surveillance. A simple Coshfibre bundle will do the job of an expensive TV camera and monitor, with no power supply and in any conditions of heat and pressure. Engineers will scrutinize the interiors of boilers and reactors; biologists will study burrowing mammals and deep-sea fauna; in shops, prisons, nurseries and public places, Big Brother will inspect his charges as never before.

More cunning still, Coshfibre can be made into a true cloak of invisibility. Each fibre will be woven so as to have its two ends facing outwards on opposite sides of the cloak. Light hitting one end will then travel along the fibre and emerge from the other end in exactly the same direction, as if it has gone straight through the cloak. Cover the entire surface of the garment with such window-ends, wrap it totally around the wearer, and light from any direction will go right through it. The cloak and its wearer will become completely invisible.

David Jones

Smallest eukaryotic organism

SIR — Autotrophic communities of picoplankton (cell size less than 2 μm diameter), known to be dominated by prokaryotes¹, are essential in the carbon cycle of estuaries² and oceans^{1,3}. Hall and Vincent report that in other nutrient-rich ecosystems, the eukaryotic forms of the picoplankton "can play a major role in generating new production"⁴. That seems to be the case for the marine Mediterranean Thau lagoon (France, 43°24' N–3°36' E) where, using flow cytometry, we have discovered a photosynthetic picoeukaryote which is the main component of the phytoplankton. This picoplankter is



a, Flow-cytometric measurement of the phytoplanktonic assemblage in Thau lagoon. Phytoplanktonic cells were distinguished according to their forward-angle light scatter (x -axis) and red fluorescence (y -axis). They were calibrated using an internal standard (0.96- μm fluorescent beads). The numbers of the new picoeukaryote (PICOE), *O. tauri*, dominate the phytoplankton community which otherwise consists of diatoms (DIAT), chlorophytes (CHLO) or cryptophytes (CRYPT). **b**, Transverse section ($\times 15,000$) of several single *O. tauri* cells. Cells generally contain no cell wall, a nucleus, one plastid with a starch granule, a very reduced cytoplasmic compartment with one mitochondrion and one Golgi apparatus.

classified as a green alga and named *Ostreococcus tauri* gen. et sp. nov. (C.C. and M.J.C.-D., manuscript submitted).

O. tauri, barely visible by epifluorescence microscopy, is detected by flow cytometry as shown in the figure (**a**), using low-forward-angle light scatter (related to cell size) and low red fluorescence (660–700 nm) due to their chlorophyll content. Electron microscopy reveals their extremely simple ultrastructure as detailed in **b**. The mean cell length and width are 0.97 ± 0.28 and 0.70 ± 0.17 μm , respectively, and the DNA content per cell is 33.31 ± 2.13 fg. *O. tauri* is thus the smallest eukaryote yet described.

Pigments are characterized by the presence of the three *a*-, *b*- and *c*-like chlorophylls, with a *b/a* ratio of 0.9 and *c/a* ratio near 0.1. The *c*-like chlorophyll was identified as the pigment Mg 3,8-divinylpheophorbide a_5 which is common to some photosynthetic eukaryotes (Prasinophyceae⁵) and prokaryotes (*Prochlorococcus marinus*⁶). Carotenoids are those of Chlorophyceae, with an unusually high violaxanthin content (3.05 fg per cell). *O. tauri*, counted bimonthly over one year, always appears as the main component of the phytoplankton in the lagoon. Cell abundances range between 10^7 and 2×10^8 cells per litre, which corresponds over the year to an average of 86 per cent of total phytoplankton cells. They are one or two orders of magnitude more abundant in Thau waters than picoeukaryote abundances reported by Hall and Vincent⁴ in upwelling areas. The annual average chlorophyll *a* biomass related to *O. tauri* is $0.51 \mu\text{g l}^{-1}$, which accounts for 28 per cent of the total biomass. In summer, when *O. tauri* is most abundant, carbon assimilation varies from 3.8 to $21.0 \text{ mg C mg(Chl } a)^{-1} \text{ h}^{-1}$. These rates are 1.3–5.4 times higher than those related to larger cells (over 2 μm diameter).

Such consistently high abundance and production raise several questions regarding the trophodynamic role of *O. tauri*. Why does this picoeukaryote species dominate the phytoplankton in the Thau lagoon? In upwelling areas⁴, or in dystrophic and acidic lakes⁷, the cold temperature and nitrate-enriched conditions, or low pH, might respectively explain picoeukaryote dominance. These factors are not relevant in temperate marine waters where the pH is stable at around 8.2. One reason for the adaptation and high production of *O. tauri* in Thau waters could be its unusual amount of violaxanthin pigment, which acts as a photoprotectant⁸. In the Thau lagoon, irradiance frequently reaches $2,000 \mu\text{E m}^{-2} \text{ s}^{-1}$ at the subsurface water layer, and the transmitted light at a depth of 4.5 m is

always more than 10 per cent of the incident radiation. Another hypothesis is related to the intensive oyster culture in the lagoon (25,000 tons produced per year). The oysters excrete large amounts of ammonium ion ($1.04 \mu\text{mol}$ in the oyster beds and $0.35 \mu\text{mol}$ outside), which favours picoplankton over the larger size classes⁹. Also, oysters preferentially retain large cells¹⁰ more than 2 μm in diameter, which could encourage the development of the cells smaller than 1 μm .

We are now examining how common this alga is, to discover whether these organisms represent a substantial but overlooked contribution to primary production in marine coastal waters.

Claude Courties

André Vaquer

Marc Troussellier

Jacques Lautier

Laboratoire d'Hydrobiologie Marine,

Université Montpellier II,

URA CNRS 1355, URM n° 5,

CC 093, 34095 Montpellier Cedex 5,

France

Marie J. Chrétiennot-Dinet

Jacques Neveux

Cordelia Machado

Observatoire Océanologique de Banyuls,

URA 117, Laboratoire Arago,

66650 Banyuls-sur-Mer, France

Hervé Claustre

Laboratoire de Physique et Chimie Marines,

La Darse, BP 8,

06230 Villefranche-sur-Mer, France

1. Stockner, J. G. & Antia, N. J. *Can. J. Fish. Aquat. Sci.* **43**, 2472–2503 (1986).
2. Malone, T. C., Ducklow, H. W., Peele, E. R. & Pike, S. E. *Mar. Ecol. Prog. Ser.* **78**, 11–22 (1991).
3. Platt, T. & Li, W. K. W. (eds) *Photosynthetic Picoplankton* (Ottawa) *Can. Bull. Fish. Aquat. Sci.* **214** (1986).
4. Hall, J. A. & Vincent, W. F. *Mar. Biol.* **106**, 465–471 (1990).
5. Ricketts, T. R. *Phytochemistry* **5**, 223–229 (1966).
6. Goericke, R. & Repeta, D. J. *Limnol. Oceanogr.* **37**, 425–433 (1992).
7. Stockner, J. G. & Shortreed, K. S. *Revue ges. Hydrobiol.* **76**, 581–601 (1991).
8. Lawlor, D. W. in *Photosynthesis* (ed. Lawlor, D. W.) 31–51 (Longman, Harlow, 1993).
9. Chisholm, S. W. in *Primary Production and Biogeochemical Cycles in the Sea* (eds Falkowski, P. G. & Woodhead, A. D.) 213–237 (Plenum, New York, 1992).
10. Desious-Paoli, J. M. in *Aquaculture, Shellfish Culture Development and Management* 319–346 (Ifremer, Paris, 1987).

Genetic dyet?

SIR — FISH (fluorescent *in situ* hybridization) has become a favourite ingredient in cytogenetic recipes. Considering the recent development in early fetal chromosome analysis, the time seems ripe to suggest that FISH & CHIPS (chromosome interphase staining) should be added to the menu.

Søren Nørby

Department of Forensic Genetics,

University of Copenhagen,

DK-2100 Copenhagen,

Denmark

HIV in infected lymph nodes

SIR — To answer questions raised by Duesberg¹, and in a separate and scientific sense by Phillips *et al.*², concerning the quantities of HIV RNA (detected by *in situ* hybridization) contained in the germinal centres of lymph nodes of HIV-infected individuals, we have constructed a test object containing a known number of HIV virus particles in a fibrin clot³. This sample can be processed in a manner identical to infected tissue specimens and examined histologically after *in situ* hybridization using a ³³P-labelled HIV antisense riboprobe cocktail representing about 9 kilobases of the HIV-1 genome⁴. An equivalent-sense probe control is used to subtract background from consecutive sections of the test object and of the tissues. Tissue specimens are digested with protease to remove proteins adherent to the viral particles⁵. The amounts of hybridization are detected by phosphorimaging the microscope slides in a Fuji BAS 2000 instrument.

Calculations show the test object to contain 15,450 viral particles per mm² of a 6-µm section. Based on a comparison of 1-mm² (6,000-µm³) portions of the test object with equal areas of lymph-node germinal centres showing maximum intensity, we estimate the number of viral particles in a germinal centre to be as much as 2.48×10^4 particles per mm². Because about 20% of the volume of lymph nodes is occupied by germinal centres in HIV-infected individuals (C.H.F., unpublished data), an entire lymph node may contain of the order of 1.2×10^9 viral particles per cm³. The amount of viral RNA appears to remain relatively constant in the germinal centres of lymphoid tissues in sequential biopsies. We have examined more than 300 tissues from individuals who had been infected for different periods of time (including gut, spleen and even lung when lymphoid aggregates occur there), which range from the time antibodies develop against HIV until up to 8 years after infection or until loss of germinal centres in the late stages of AIDS. A similar circumstance occurs in other primates infected with SIV.

In reply to questions regarding the infectious nature of such viruses, we refer readers to a series of relevant papers by the late Albert Sabin⁶, in which he re-

ported experiments showing that viruses that had been reacted with antibodies retain their infectivity in tissues. It seems reasonable to us that CD4-expressing T lymphocytes become infected following collision with stored, protein-cloaked virus, as the cells traffic through lymphoid germinal centres. There may also be loss of CD4 expression when infected T cells express HIV. Projected over years, the rate of T-cell replacement falls behind their loss and, combined with alterations of cell populations and the developmental microenvironment of the germinal centre arising from the infection, the balance of both structure and function in the immune system is catastrophically altered. We believe that the protein-cloaked virus in germinal centres, combined with some

element of follicular dendritic cell function⁷ and T-cell kinetics, is the explanation for the extremely slow but inexorable progression of the primate lentiviral diseases.

Cecil H. Fox

Shelley Hoover

Victoria R. Currall

*Molecular Histology Laboratory, Inc.,
7605 F Airpark Road,
Gaithersburg, Maryland 20879, USA*

H. James Bahre

*Fuji Medical Systems,
Stamford, Connecticut 06912, USA*

M. Cottler-Fox

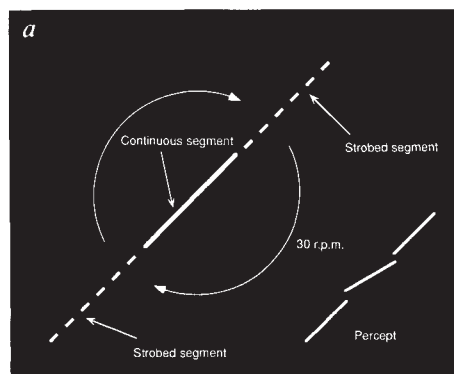
*Department of Transfusion Medicine,
Warren G. Magnuson Clinical Center,
National Institutes of Health,
Bethesda, Maryland 20892, USA*

Motion extrapolation in catching

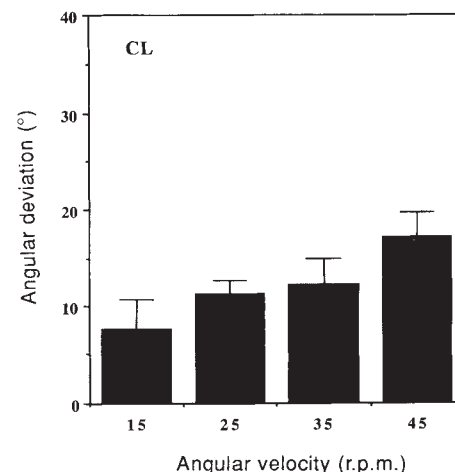
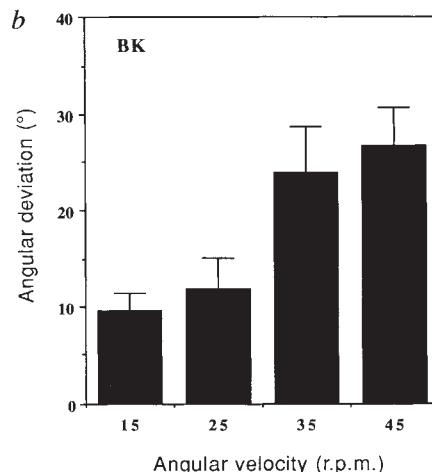
SIR — Many studies investigate how observers might compensate for the latency due to bodily action while executing behaviours such as catching a ball. But success in such interceptive behaviour depends on the observer having accurate information about the ball's initial location. This raises a question that has received surprisingly little attention. How does the visual system compensate for the delay in the transmission of motion information from photoreceptors to 'higher' visual areas of the brain? The problem is

serious because the typically estimated delay (τ) of about 100 ms (ref. 1) would cause an object moving at 30 m.p.h. to appear retarded by 4.4 feet in its trajectory!

An effect first observed by MacKay² and later rediscovered in our laboratory³ in the following form, suggests an answer. A single white line is rotated against a black background (see *a* in the figure). The line consists of one continuously illuminated segment and two strobed segments. Observers report a compelling



a, The stimulus was a single physical slit (3.9° long) rotating continuously at 30 r.p.m. against a dark background. The central 1.3° of the slit was continuously illuminated (solid line) and the two outer 1.3° segments were strobed (dashed lines) for 5 ms. Ten observers (six naive) reported a compelling spatial lag of the strobed segments (percept). *b*, The spatial lag increased as a function of the angular velocity of the slit. This lag was measured by having two observers (BK and CL) turn a disk to rotate the strobed segments relative to the continuous segment until they appeared to neither lag nor lead.



1. Duesberg, P. *Biotechnology* **11**, 421 (1993).
2. Phillips, A. N. *et al.* *Nature* **367**, 124 (1994).
3. Cottler-Fox, M. & Fox, C. H. *J. infect. Dis.* **164**, 1239–1240 (1991).
4. Fox, C. H. & Cottler-Fox, M. *Curr. Protocols Immun.* **12**, 1–21 (1993).
5. Fox, C. H. *et al.* *J. infect. Dis.* **164**, 1051–1057 (1991).
6. Sabin, A. *Br. J. exp. Path.* **16**, 70–101; 158–180 (1935).
7. Fox, C. H. & Cottler-Fox, M. *Immun. Today* **13**, 353–356 (1992).

spatial lag of the strobed segments, which increases with the angular velocity of the line (b in the figure).

How might this effect be explained? I believe the answer lies in the predictability of the continuous segment and the unpredictability of the strobed segments resulting in a differential processing delay. The continuous segment is illuminated for a long enough time for it to be represented cortically. Furthermore, as the angular velocity of this segment is constant, there is information at time t as to where the segment will be at $t + \tau$ ms. It is proposed that in order to overcome the transmission delay, an 'early' visual mechanism corrects the spatial lag by extrapolating the moving object's instantaneous location. Thus, the perceived location, which incorporates the input from this mechanism, is closer to the object's physical location than might be expected from neurophysiological estimates¹ of the delay. The perception of the strobed segments is also contingent on the retinal signal triggering a cortical neural representation, but owing to the unpredictability of the stroboscopic event, the visual system cannot overcome the transmission delay in this case.

If an equal delay for both the moving and the strobed segments is assumed, one could account for the present 'flash-lag' effect in terms of visual persistence of the strobed segments for about 100 ms after their off-set, and the 'deblurring' of the continuous segment by the motion system^{3,4}. But according to this account, the strobed segments should appear aligned with the continuous segment at the instant of strobe onset. While observers report that the large visible misalignment (of up to 25°) is already present at the instant of strobe onset.

The present findings suggest that in the case of moving objects the visual system overcomes at least some of the transmission latency through extrapolation. Future experiments will reveal the discrepancy, if any, between the extrapolated and physical locations of moving objects. In this context, it is interesting to note that the computed average (angular deviation ÷ angular velocity) of the time delay of the strobed segments is approximately 82 ms, which is not too different from typical estimates of about 100 ms. Thus, the error between the physical location and the extrapolated location may not be very large.

Romi Nijhawan

Department of Psychology,
Uris Hall, Cornell University,
Ithaca, New York 14853, USA

Ant sex ratios

SIR — Sundström¹ reported that colonies of the ant *Formica truncorum* on islands in southern Finland produce sexual broods with a bimodal distribution of sex ratios. Colonies headed by queens that Sundström inferred had mated multiply (based on intra-colony allele frequencies at 3–4 allozyme loci) rear significantly more males than females, indeed sometimes only males, whereas colonies headed by putatively monogamous queens do the reverse. Sundström argued that sex specializations result from workers first assessing their relatedness to male versus female brood, relative to average worker-relatedness asymmetries in other colonies of their population, and then adjusting the sex ratio in their own reproductive interests (specializing on females when relative relatedness asymmetry is high, and on males when it is low).

The hypothesis of genetic-relatedness asymmetry^{2,3} requires that worker ants: (1) assess the total number of different males with whom their mother mated (which conceivably exceeds the number of patrines active during any one worker's lifetime) relative to the mean frequency of polyandry in their local population; and (2) recognize the sex of eggs or larvae, and behaviourally bias their colony's sex ratio, for example via neglect or siblicide; but (3) do not distinguish full sisters from half sisters among eggs or larvae (because discriminative nepotism would always result in female-biased broods). Because Sundström^{1,4} provided no evidence that *F. truncorum* workers are altering their queen's preferred sex ratio, it is not possible to judge the plausibility of the mechanisms underlying her interpretation.

An alternative hypothesis to explain bimodal sex ratios in ants² is that queens themselves sometimes lay predominantly male (haploid) or female (diploid) eggs, with the workers caring for all the brood. This hypothesis implies synonymy of queen and worker sex-ratio preferences, a possibility that has apparently not been investigated for any species. A direct test would be to compare primary versus secondary sex ratios — that is, samples of reproductive-destined eggs versus alates, controlling for workers' elimination of inviable eggs.

Queens and workers would benefit from manipulating the sex ratio in that outbreeding would be optimized. Colonies of social insects are frequent targets of debilitating parasites and pathogens^{5,6}, and dispersal from infected areas and genetic variability among progeny might thwart such biotic enemies. Multiple queens, multiple mating by queens and outbreeding all enhance genetic variability^{5,6}. Male-biased broods promote both extreme dispersal and out-

breeding, especially since male ants are typically much smaller and lighter than reproductive females.

We suggest that queens adjust their mating frequency and, in collaboration with workers, their progeny sex ratio in response to the local severity of parasites and pathogens. If this hypothesis were correct for Finnish *F. truncorum*, queens from islands where diseases are prevalent, where there are local pockets of disease within islands, or where there are survivors from individual diseased colonies, would mate multiply and produce male-biased, dispersive broods. Queens from relatively disease-free localities (for example where the success of daughter colonies is high), in contrast, would mate monogamously and produce female-biased, philopatric broods. This suggestion is an alternative explanation for Sundström's¹ fascinating results.

Paul W. Sherman

Janet S. Shellman-Reeve

Section of Neurobiology and Behavior,
Cornell University,
Ithaca, New York 14853, USA

SUNDSTRÖM REPLIES — Sherman and Shellman-Reeve offer constructive suggestions for further investigation, together with some new ideas on causes of sex ratio specialization in ants. I agree that queens may gain at least partial control by adjusting the primary sex ratio, and this is clearly the next question to address. In colonies headed by a multiply mated queen, they propose that workers and queen mutually agree to produce a male-biased sex ratio. This is also fully consistent with worker control according to current sex ratio theory, because the optimal sex ratio for workers and queen converge in colonies headed by a multiply mated queen^{7,8}.

However, for the colonies headed by a singly mated queen, Sherman and Shellman-Reeve make some predictions which are not supported by the data. First, if females disperse less than males, as they suggest, local resource competition among females would ensue, with selection for male-biased population-level sex ratios⁹. This stands in contrast to my results, which demonstrate a female-biased population-level sex ratio. Moreover, under local resource competition only small colonies may be expected

1. DeValois, R. L. & DeValois, K. K. *Vision Res.* **31**, 1619–1626 (1991).

2. MacKay, D. M. *Nature* **181**, 507–508 (1958).

3. Nijhawan, R. *Invest. Ophthalmol. vis. Sci., Suppl.* **33**, 1415 (1992).

4. Burr, D. *Nature* **284**, 164–165 (1980).

1. Sundström, L. *Nature* **367**, 266–268 (1994).

2. Boomsma, J. J. & Grafen, A. *Evolution* **44**, 1026–1034 (1990).

3. Queller, D. C. *Am. Nat.* **142**, 346–351 (1993).

4. Sundström, L. *Behav. ecol. Sociobiol.* **33**, 345–354 (1993).

5. Sherman, P. W., Seeley, T. D. & Reeve, H. K. *Am. Nat.* **131**, 602–610 (1988).

6. Hamilton, W. D. in *The Natural History of Inbreeding and Outbreeding* (ed. Thornhill, N. W.) 429–450 (Univ. Chicago Press, 1993).

7. Boomsma, J. J. & Grafen, A. *J. evol. Biol.* **3**, 383–407 (1991).

8. Trivers, R. L. & Hare, H. *Science* **191**, 249–263 (1976).

9. Frank, S. A. *Behav. Ecol. Sociobiol.* **20**, 95–201 (1987).

to produce an all-female brood, which again is contrary to my findings. The model thus fails to explain the female-biased population-level sex ratios as well as the occurrence of all-female broods over the whole range of colony sizes. Hence, in *Formica truncorum*, the female-biased population-level sex ratios, together with the colony-level sex ratio patterns, still support predictions based on worker control of sex allocation^{2,7,8}. The effect of parasites/pathogens on sex allocation is, however, certainly a factor that needs to be addressed, especially as multiple mating may confer higher disease resistance⁵.

Liselotte Sundström*

Bern University, Ethologische Station Hasli,
Wohlenstrasse 50a, Hinterkappelen,
Switzerland

*Also at University of Lausanne, BB, Institut de Zoologie et d'Ecologie Animale, 1015 Lausanne, Switzerland.

Man-made lakes and sea-level rise

SIR — Sahagian *et al.*¹ estimate the total man-made lake impoundment to date to be $1.9 \times 10^{12} \text{ m}^3$. As such it is the single largest (and negative) contribution to the sea level, yet it is not large enough to offset the combined positive contributions from other anthropogenic sources such as groundwater withdrawal and deforestation. However, my earlier estimate², based on documented sources^{3,4}, puts the total water impoundment (by 1991) at as much as $10 \times 10^{12} \text{ m}^3$. More than five times larger than that of Sahagian *et al.*¹, this estimate would easily swamp all the other positive contributions (as I explicitly stated²) and the sum thereof.

Just for the sake of argument, suppose that on average the reservoirs are only filled to three-quarters of their designed capacity, and that no new reservoirs were constructed after 1986 (which is not in fact the case, considering some gigantic water projects in developing countries such as China). Then, by proper scaling of my earlier estimate, the total water impound-

ment would be about enough to nullify the other anthropogenic sources, leaving a zero net contribution.

Thus, it is clear that man-made water impoundment is the single element that decides whether the total anthropogenic contribution to the global sea level is positive (as claimed in ref. 1), negative (as claimed in ref. 2), or close to nil (as in the above hypothetical example). To resolve this only requires some up-to-date compilation of basic data for the world's water projects, a task much more accessible than natural processes such as that of calculating water balance of mountain glaciers or polar ice sheets.

Benjamin Fong Chao

Geodynamics Branch,
NASA Goddard Space Flight Center,
Greenbelt, Maryland 20771, USA

SIR — Sahagian *et al.* attempt to estimate the share of observed sea-level rise that can be attributed to human activities rather than global warming (and consequent glacial melting and thermal expansion of sea water)¹. Unfortunately, the authors significantly understate the amount of fresh water stored behind dams, an error that when corrected cancels or even reverses the claimed effect.

To estimate a net sea-level rise of 11.8 mm from non-climate-related sources during the twentieth century, Sahagian *et al.* summed total contributions from several anthropogenic sources (for example, depletion of 'fossil' aquifers, loss of soil moisture and deforestation). They then deducted the amount of water stored behind dams and thus removed from the sea-level equation during this period. They reported this storage to be $1.9 \times 10^{12} \text{ m}^3$, the equivalent of a 5.2-mm drop in sea level.

The 1983 source⁵ used by Sahagian *et al.*, however, actually listed 170 major dams with a combined reservoir capacity of $2.8 \times 10^{12} \text{ m}^3$, $0.9 \times 10^{12} \text{ m}^3$ greater than they used in their calculations. And by 1986, there were at least 36,065 additional reservoirs behind large dams that are 15 m tall or over (an additional 2,000 might have been unreported from the former Soviet Union)⁶. There are literally millions of smaller reservoirs, down to the size of farm ponds, all over the world.

Total reservoir storage has been variously estimated at $5.5 \times 10^{12} \text{ m}^3$ (in 1983)⁷ and more than $5.0 \times 10^{12} \text{ m}^3$ (in the late 1980s)⁸. Even one source¹ cited in ref. 9 estimated that total reservoir volume (large and small) grew by $6.0 \times 10^{12} \text{ m}^3$ between 1932 and 1982 and that total anthropogenic storage during the same time, including recharging of ground water and increased soil moisture due to irrigation projects, grew by $18.8 \times 10^{12} \text{ m}^3$ — translating to a 5.2-cm sea-level decrease¹⁰, ten times larger than the figure used in ref. 1. Thus, depending on the

figures used, net sea-level rise from anthropogenic sources during this century ranges from almost nil (a mere 3.3 mm) to strongly negative (−35.0 mm).

Correcting the analysis in ref. 1 makes it clear that little, if any, of the global sea-level rise we observe can be assigned to the anthropogenic causes listed therein. Indeed anthropogenic storage of water behind dams, in aquifers and as soil moisture, could well have significantly limited the observed sea-level rise. There is no reason to think that the contributions of climate-related effects are smaller than previously supposed.

Eric Rodenburg

World Resources Institute,
1709 New York Avenue,
Washington, DC 20006, USA

SAHAGIAN ET AL. REPLY — Chao and Rodenburg both suggest that there is more water stored in artificial reservoirs than we accounted for in our paper¹, and that this would reduce the amount of net anthropogenic sea-level rise to date. We do not doubt that our list of artificial reservoirs was incomplete, as we did not include dams listed as under construction⁵.

In assessing the consequences of this omission, there are two important issues to consider. The first is the difference between reservoir capacity and the actual amount of water stored in the reservoir. In our analysis we assumed that all of the dammed reservoirs have been continuously filled to capacity (and free of sediments). This is surely not the case. Chao's supposition of three-quarter capacity may be reasonable, but such data have not yet been compiled.

Second, the surge of large dam-building activities between 1954 and 1985 has essentially ceased, and other than the Three Gorges Dam in China, there seem to be no major future plans for large dams nor is there any reason to believe that the number or volume of small farm ponds, as suggested by Rodenburg, will increase.

Consequently, we conclude that while additional dammed reservoirs not included in our tally would reduce the calculated historical amount of anthropogenic sea-level rise, they do not affect the present rate of anthropogenic sea-level rise (to be extrapolated into the future). It is the present rate of sea level rise, and not the past amount, that will have the greatest impact on modern society. To assume that the continued construction of large dams will abate future anthropogenic contributions to sea-level rise is not, in our view, wise.

D. L. Sahagian

D. K. Jacobs

F. W. Schwartz

Byrd Polar Research Centre,
Ohio State University,
Columbus, Ohio 43210, USA

1. Sahagian, D. L., Schwartz, F. W. & Jacobs, D. K. *Nature* **367**, 54–57 (1994).
2. Chao, B. F. *EOS Trans. Am. Geophys. Un.* **72**, 492 (1991).
3. Chao, B. F. *J. geophys. Res.* **93**, 13811–13819 (1988).
4. UNESCO *World Water Balance and Water Resources of the Earth* (UNESCO Press, USSR Committee for the International Hydrological Decade, 1978).
5. Tilsley, J. M. *Major Dams, Reservoirs, and Hydroelectric Plants*. (US Department of Interior, US Bureau of Reclamation, Denver, 1983).
6. International Commission on Large Dams *World Register of Dams: 1988 Updating* (International Commission on Large Dams, Paris, 1988).
7. Golubev, G. N. *Hydrol. Sci. J.* **28**, 57–75 (1983).
8. Shiklomanov, I. A. in *Water in Crisis* (ed. Gleick, P. H.) 13–24 (Oxford Univ. Press, New York, 1993).
9. Newman, W. S. & Fairbridge, R. W. *Nature* **320**, 319–321 (1986).
10. Warrick R. & Oerlemans, J. in *Climate Change: The IPCC Scientific Assessment* (eds Houghton, J., Jenkins, G. & Ephraums, J.) 257–281 (Cambridge Univ. Press, 1990).

Not seeing eye to eye

C. R. Cavanaugh

In the Eye's Mind: Vision and the Helmholtz–Hering Controversy. By R. Steven Turner. Princeton University Press: 1994. Pp. 338. \$49.50, £37.50.

MODERN science is increasingly done by committee: credit and blame are spread over many authors, whose publications are a bland lowest-common-denominator of their opinions. In the nineteenth century, single authors were the norm and the level of their polemic can seem strange to modern eyes. This personal quality had other consequences: even the most eminent scientists tended to make their own observations; and the higher one's standing, the more credence the results received. In the hands of a careful, shrewd observer this could be beneficial, but it could also be disastrous, as when Blondlot deluded himself, the French Academy of Sciences and many respected scientists into believing in the existence of N-rays; when shown to be in error, he went mad.

Steven Turner gives a scholarly account of two of the most famous antagonists of that era: Hermann von Helmholtz and Ewald Hering. Of the two, Helmholtz is better known today, largely because of his remarkable ability to ride the crest of the wave of contemporary science. Hering was in comparison an outsider, but also more of a visionary: at a time when the very existence of single neurons was disputed, he proposed what are today called Hebbian synapses, whose strength is modified by use.

Although Hering is largely forgotten, a century ago he was more nearly Helmholtz's equal: in the edition of Poggendorff's biographical dictionary that covers the period 1904–22, it is Hering who is given more space (although this must have reflected his then recent death). Hering's entry in 1958 is similar to that in the earlier edition, whereas Helmholtz's has grown by an astonishing 50-fold.

Hering tends to sound irascible in his writings, but his modern characterization as a solitary misanthrope cannot be accurate. He was clearly capable of negotiation and persuasion, since he was a leading force in establishing, and was the founding rector of, the German University in Prague, which was for half-a-century a centre of academic excellence in central Europe. Further, he gathered around him a group of devoted followers, which Helmholtz never achieved.

Although they disagreed on many topics, colour vision is perhaps the most accessible and certainly the best-known. Helmholtz followed Thomas Young in holding that there are three types of receptor, each responding to a limited portion of the visible spectrum such as

red, green or blue, and that all other colour sensations arise when these receptors are stimulated in various ratios. Hering also believed in three types of receptor (although these were not necessarily the retinal cones), but postulated that activity in each type is increased by light of one colour and decreased by its complementary colour. This resulted in three opponent processes, one sensitive to variation along a red–green dimension, one along yellow–blue and one along black–white. This proposal, which was fundamentally qualitative and supported by such observations as the fact that one can see reddish yellows or greenish blues, but not reddish greens or yellowish blues, found disfavour among the chemists, who could find no photochemical reactions with this property, and among the physiologists, who held that stimulation could increase, but not reduce, nervous activity.

Interestingly, neither scheme originated with the person with whom it is

usually associated. Twenty years before Young, a London merchant named George Palmer proposed a theory of colour vision that is essentially identical to Young's. Palmer's ideas were described in Lichtenberg's *Magazin*, a periodical that had much the same function as the News section of *Nature* today and with which Young was familiar. Curiously, although Turner acknowledges the advice of J. D. Mollon, who has written on Palmer, he follows tradition in attributing the trichromatic theory to Young. Hering was anticipated by Arthur Schopenhauer, who in his youth was much influenced by Goethe and who in 1816 proposed an explicit opponent-process model. His book went into two further editions, the last published in 1870, when Hering was formulating his ideas on vision. Although the opponent colours that Schopenhauer proposed (red–green, orange–blue and yellow–violet) differ from Hering's, the principle is similar; and considering Schopenhauer's prominence, Hering must have been aware of his theory.

In the course of time, the balance tipped firmly in favour of Helmholtz, as colour science became increasingly identified with colorimetry, the branch of applied physics that deals with specifying colours in terms of mixtures of certain real (or better, ideal) primary colours. Because



The Magic Lantern by Charles-Amédée-Philippe van Loo (1764). Taken from Barbara Maria Stafford's *Artful Science: Enlightenment Entertainment and the Eclipse of Visual Education*. MIT Press, \$35, £24.95.

Of famous London town



"DIRTY odious London," Darwin called it, but to the modern scientific tourist this great city can be paradise. The newly published *London Science* (pp. 224, Prion, £14.99) by Dennis and Sylvia Rosen is sure to become the essential Baedeker to anything in the capital of scientific, technological and medical interest.

Its colourful chapters cover "Repositories of science", with information on more than 60 museums, including lesser-known specialist collections such as the British Optical Museum and the Vintage Wireless Museum, as well as more familiar ones such as the Imperial War Museum (above, left) and the Natural History Museum (above, top right); "Guar-

dians of the sciences" (teaching centres and learned societies); "Where the work is done" (libraries and laboratories); "Works of art", which looks at science and scientists through artists' eyes; "Memorials to the practitioners" (Alexander Fleming, for example, has a public house named after him (above, bottom right)); "Shopping for science" (instruments, books, antiques and children's science); and "Journeying with the sciences", which describes four walks and an excursion that takes in, among other sights, the Thames Barrier, one of Europe's most remarkable engineering achievements (below). Fascinating reading for inveterate travellers and Darwinite recluses alike. P.T.



only three primaries are needed to specify any colour, and because nothing analogous to Hering's opponent processes is needed, it was widely assumed that colour vision could be accounted for by Young's trichromatic system. By 1950, textbooks cited the trichromatic model as *the* theory of colour vision and, with the notable exception of Leo Hurvich and Dorothea Jameson, who almost single-handedly championed opponent processes, Hering was viewed as an historical aberration. But while the trichromatic model can explain why mixtures of colours that are physically different can appear the same to our eye, it cannot explain many facts about the appearance of colours — as Hering and his followers had frequently pointed out and Helmholtz tacitly acknowledged.

At about this time, a remarkably sudden change occurred, in part as a result of demonstrations that single neurons in the visual system can be excited by certain wavelengths and inhibited by others. Although the resulting data were so ragged as to have been derided had they arisen from psychophysical experiments, electrophysiology is inexplicably viewed as objective in a way that psychophysics is not. Colour theories soon became variations on the 'zone' theories of the turn of the century, in which the input level contains three types of Helmholtzian receptor, the outputs of which are combined at the next level in various positive and negative combinations so as to create opponent processes of the Hering type.

Turner has clearly done his homework and — unlike many people who write on the history of colour vision — has read the original texts. He understands the issues and the methods used in studying them, and does an excellent job of defining the jargon of the era, which is often comprehensible only in context. I would personally have liked to have read more about the actors in this drama: as it is, they scarcely seem to have been living persons. And Turner's scholarly approach may miss the point of why neither Helmholtz nor Hering could accept the zone theory that Johannes von Kries and others offered them. Turner proposes an explanation in terms of Kuhnian "perceptual incommensurability". I suspect the real reason was simpler: sheer bloody-mindedness. The two men disliked each other profoundly, and neither would accept a compromise.

It is unfortunate that an otherwise well-produced book has fallen into the clutches of that desktop publishing demon, the lost data-file: all bibliographic references to Helmholtz are missing. The publisher should try to make them available. □

C. R. Cavonius is in the Institut für Arbeitsphysiologie, University of Dortmund, Ardeystrasse 67, D-44139 Dortmund, Germany.

Rocky motoring

Euan Nisbet

Understanding Earth. By Frank Press and Raymond Siever. *W. H. Freeman*: 1994. Pp. 593. \$44.95, £23.95 (pbk).

Exploring Earth and Life Through Time. By Steven M. Stanley. *W. H. Freeman*: 1993. Pp. 538. \$45.95, £22.95 (pbk).

A History of the Earth. By John J. W. Rogers. *Cambridge University Press*: 1994. Pp. 312. £50, \$89.95 (hbk); \$39.95, £22.95 (pbk).

A GOOD text can apologize for a hundred bad lectures: North America's generally excellent system of broadly based 100-level university classes depends heavily on superb introductory texts that more than compensate for the drone of the mind-dead professor such as myself, plodding away while the hot-shots get on with their grant proposals. One of the curious rituals of spring in academia is the flight of the condors down the dusty corridors of the science departments, when powerful and elegant publishing creatures, wafting Chanel, turn their attentions to these tired professors. They swoop down, clutching the latest blockbuster text, and stun their victims into signing up for hundreds of copies. As the former teacher of a 'Rocks for Jocks' class of 400 nonscience students, including a good hunk of the football and ice-hockey teams, as well as teachers, nurses and nuns, I used to be prime meat for the sales reps. When one was spotted winging raptorially down the corridor, it was time to go into hiding for the morning, as I preferred to make my decisions in the quiet of a good bookshop. Despite the sales drives, for my first-year nonscience students there were surprisingly few good choices of text, and in the end I wrote my own.

Two new books, 'downmarket' versions of well-known works by Frank Press and Raymond Siever and by Steven Stanley, should fill this important and neglected niche. The texts almost make it worth awaiting the circling reps, who will no doubt be hawking them about vigorously.

As expected from the record of the authors, these are fine books, well illustrated and thoroughly thought out. The books do not have the flair of a good single-author text, bubbling with ideas and interest. Instead, the authors and their production teams have opted for the detailed solid presentation and well-worked contents that will get the student from A to B no matter how badly the class is taught. The books are not Ferraris, but reliable, roomy and comfortable Chrysler minivans that can run safely on cruise control with only minor correction.

The texts differ strongly in emphasis. Press and Siever's work is strictly on

physical geology, a steady drive through the introductory chemistry, physics, tectonics, petrology and geomorphology of our planet. By contrast, Stanley takes a more biological route. Starting with a discussion of time and environment, he moves to the rock record and a brief excursion into tectonics, followed by a tour through the geological column, admiring its varied inhabitants from bacteria to dinosaurs. Both books are careful, clearly written and mostly reliable. I have my complaints, of course. The content of both books is distinctly elderly, reflecting the conventional wisdom of the 1970s, though to be fair there are frequent glimpses into modernity. The ageing is most serious in the discussion of subduction in Stanley's book, which is in need of major updating. Press and Siever also have their weaknesses, for instance in the treatment of the formation of oceanic crust at mid-ocean ridges. Given the importance of topics such as these, great care is needed in handling them, as it is often difficult to persuade students that a textbook is wrong.

Press and Siever's title, *Understanding Earth* (is it understanding? — I've thought of it as a tough old rogue, given to the shakes) is not to be confused with the excellent and newly reincarnated book *Understanding the Earth* edited by G. C. Brown, C. J. Hawkesworth and R. C. L. Wilson (Cambridge University Press, 1992), but the book by John Rogers might well be, as it does try to comprehend the planet. Rogers is a distinguished Earth scientist, and this book reflects an erudite and intelligent distillation of the contemporary view of our planet. Here is a book for advanced students that is fun to read, provokes thought, interest and annoyance, is not glossy (some of the diagrams are engagingly lousy) and has an integrity of argument and purpose that should make it perfect for reading in the washroom while one hides from the publishing representatives. The discussion is provocative and thoughtful, generally interesting, contemporary and reliable (though as a southern African whose family has for 150 years puzzled over the geological links between the Cape and the Falkland Islands, I must protest at the extreme political incorrectness of the "Malvino-kaffric" realm!). The bias is tectonic (unlike Preston Cloud's *Oasis in Space* (Norton, 1988) which is mainly biological), but there is also plenty of biology. It is well worth delving into, or raiding for lecture material. Although not a Ferrari, the book is perhaps best described as a Volkswagen with a V8 engine installed — powerful and enjoyable transportation for a higher-year core course. □

Euan Nisbet teaches in the Department of Geology, Royal Holloway, University of London, Egham, Surrey TW20 OEX, UK.

Ancients and moderns

George C. McGavin

The Insects: An Outline of Entomology. By Penny J. Gullan and Peter S. Cranston. *Chapman and Hall*: 1994. Pp. 491. £24.99, \$39.95 (pbk).

Latin American Insects and Entomology. By Charles L. Hogue. *University of California Press*: 1993. Pp. 536. \$85, £60.

The Bee Genera of North and Central America. By Charles D. Michener, Ronald J. McGinley and Bryan N. Danforth. *Smithsonian Institution Press*: 1994. Pp. 209. \$53.95, £34.95.

Quaternary Insects and their Environments. By Scott A. Elias. *Smithsonian Institution Press*: 1994. Pp. 284. \$47.95, £31.25.

SINCE receiving *The Insects* I have referred to it daily, my other standby reference texts meriting only an occasional glance. It is an extremely well written and thoroughly up-to-date text, amply and clearly illustrated with magnificent black-and-white drawings. The authors have adopted a systems-based approach: there are large sections, for instance, on sensory systems and behaviour, aquatic insects, insects and plants, insect societies and pest management. Basic information about each insect order is provided in separate boxes, which are also used to deal with specific topics of interest such as figs and fig wasps, insect binary chemical weapons and biological clocks. Students wanting a single book to give them the whole picture need look no further. This work will appeal to those with only a passing interest in insects as well as rapidly become the standard teaching textbook on entomology.

From the crowded canopies of the rainforests and the vastness of the grasslands to the wilderness of the high Andes, there can be little doubt that South America is one of the most exciting parts of the world. Whether you are an armchair naturalist or a machete-wielding field biologist, you will find much that educates and enchants in *Latin American Insects and Entomology*. Despite the title, the text also covers selected terrestrial crustacean, arachnid and myriapod families.

The first fifth of the work is introductory, covering the history of Latin American entomology, general insect biology, South American ecology and economic entomology. The rest of the book reviews selected arthropod taxa (orders, families and some species), giving Spanish, Portuguese and English names, covers aspects of their biology and lists key references. It is hard to categorize this quirky but readable book. What the author has attempted and, to my mind, succeeded in doing is to

convey his immense enthusiasm and life-time's neotropical experience. The text is exciting and accessible by not being overly technical and detailed. This is not a convenient pocket-sized volume but a veritable mine of information that will more than earn its space in one's luggage.

Bees are among the most important insects. By the age of twelve a Yanomamö boy can recognize 12 species of stingless bee (Meliponinae) from their structure, behaviour and nest architecture. Armed with a binocular microscope and *The Bee Genera of North and Central America*, even a novice will now be able to identify any of the 169 bee genera found north of the Colombia–Panama border. If only there were such a practical information source and identification guide for every major insect group and faunal region. The book is authoritatively written and illustrated by clear line drawings, photographs and scanning electron micrographs. The excellent, skilfully constructed key, which fills nearly half of the book and appears with parallel Spanish text, works backwards as well as forwards. To make identification even easier and the key even more user-friendly, if a specimen agrees with one of seven locators (based on fundamental characters), one is able to bypass large portions of the key text at a stroke. The book also includes information on collecting and preserving specimens and clear and concise notes on each genus. The authors directly address biodiversity issues and the specific needs of entomologists, ecologists and pollination botanists.

Quaternary Insects and their Environments reviews, for the first time and in a lucid and scholarly manner, the history, use and significance of fossil insect remains in the study of palaeoecology, palaeoclimatology, zoogeography and archaeology. Until now this information has been scattered throughout specialist journals. Insects respond rapidly to changes in the environment and it is this that gives entomologists a unique insight into the recent past. Insects can be sensitive and accurate bioindicators and therefore tell us a lot about climate, the environment and human activity during the past 1.8 million years. Every biologist, geologist or environmental scientist should contemplate getting involved in this exciting and rapidly developing field of research. I, for one, look forward to the opportunity of exploring some organic deposits with a trowel and a sieve. □

George C. McGavin is at the Oxford University Museum, Parks Road, Oxford OX1 3PW, UK.

Addendum

Moon Shot, reviewed by Arthur C. Clarke in *Nature* **370**, 107 (1994), is published in the United Kingdom by Virgin at £15.99.

What we ate

Peter D. Moore

Tropical Archaeobotany. Edited by Jon G. Hather. Routledge: 1994. Pp. 270. £45, \$65.

MAN may not be able to live by bread alone, but it certainly helps. The exploitation of plant products for human consumption has had a vital role in the development of civilization, but most archaeological work on this topic has been centred on the temperate regions. The moist and monsoonal tropics of Africa, the Indo-Pacific and Latin America form the geographical focus of this collection of papers that arose from an archaeological conference held in Venezuela in 1990.

The volume admirably illustrates the impressive range of techniques now available to archaeologists attempting to reconstruct prehistoric agricultural practices and early human diets. Charcoal and carbonized plant materials survive well in archaeological sites and can result from the intentional burning of fuel wood and refuse, or the unintentional combustion of food. Both sources can offer insights into human culture. Charred roots and tubers, for example, have been identified by anatomical analysis and have permitted detailed reconstruction of early subsistence farming practices in the Pacific region.

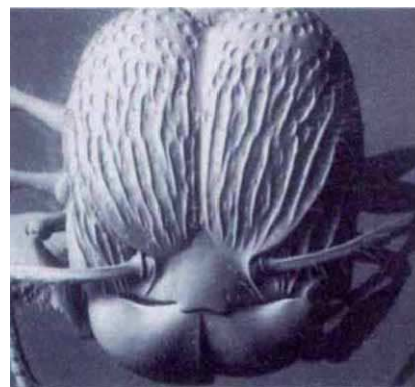
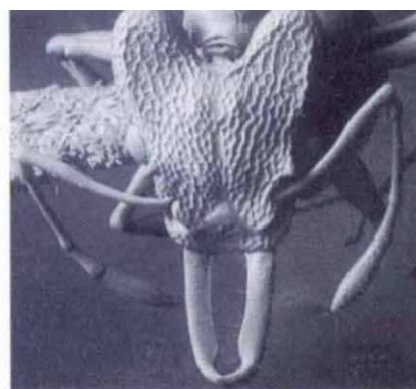
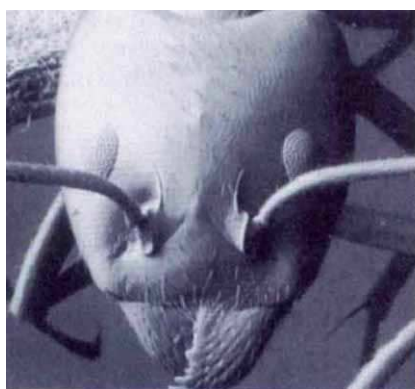
Human gut contents provide direct evidence of past tropical diets, especially

where desiccation has resulted in good preservation, but only materials that are resistant to digestion and post-mortem decay survive. Soft tissues are unfortunately lost but, even so, gut contents have fewer interpretative problems than, say, coprolites. The analysis and identification of starch residues on stone tools provide a further technique for dietary determination that avoids the effects of digestion.

Among microfossils, phytoliths are proving valuable not only for the study of grasses but also for many other plant families, from palms to gourds. Fossil pollen can also provide direct evidence of cultivation where the crop species is identifiable (as in the case of *Cocos nucifera*, but probably not in the case of cereals) or it can act as an indicator of vegetation disturbance when weed communities or the replacement of forest by grassland are detected. Perhaps the most novel techniques described here, however, relate to the use of biochemical analysis of small fragments of plant material that may permit identification of food species. Analysis of protein, lipid and DNA promise much for future studies.

The diverse nature of plant exploitation in the wet tropics is becoming increasingly apparent with the widespread application of these techniques. As far as early tropical agriculture is concerned, man evidently did not live by bread alone. □

Peter D. Moore is in the Division of Life Sciences, King's College, Campden Hill Road, London W8 7AH, UK.



ALL myrmecologists — according to Edward O. Wilson there are no more than 500 in the world — will surely want to own a copy of Barry Bolton's *Identification Guide to the Ant Genera of the World* (Harvard University Press, \$65, £51.95). As well as hundreds of scanning electron micrographs that illustrate its taxonomic keys, this comprehensive volume contains a reference list, glossary of morphological terms, and notes on broad distribution and specimen preservation.

Dynamics of chemical processes in polar solvents

Peter J. Rossky & John D. Simon

Understanding solution-phase chemistry requires a microscopic description of the electronic structure of the reacting molecules, and of the complex influence of the solvent medium on the reaction energetics and dynamics. Processes involving reactant charge-transfer are of particular importance in chemistry and biochemistry, and are strongly influenced by polar media. Recent advances in experimental ultrafast laser spectroscopy and in computer simulation are working together to provide insight into the underlying molecular principles governing this class of processes in solution.

MOST chemical reactions take place in solution. This simple fact has spurred a tremendous research effort to elucidate how solvents affect chemical reactivity. A complete understanding of solution-phase chemistry requires knowledge of the detailed molecular motions that are involved in the reaction dynamics. The elucidation of a set of unifying principles describing chemical dynamics in liquids presents challenges to both theoreticians and experimentalists. The amorphous and dynamic nature of the liquid state leads to a reaction system of exceptional complexity.

To see the origin of this complexity, one can consider first the simplest description of a solvent, one that relies only on bulk properties (density, viscosity and dielectric response). Compared to the reaction in the gas phase, bulk properties of the solvent alone can be expected to alter both the equilibrium potential-energy surface, including the activation free-energy barrier (ΔG^*), and the overall free-energy change (ΔG_0). This is a result of the different effects of solvation on the reactants, products and the transition state. Quantifying these effects on the basis of bulk solvent properties has led to the development of free-energy relationships between ΔG^* and ΔG_0 (ref. 1).

But we know that in the gas phase the act of crossing the reaction barrier can be a very fast event, occurring on the femtosecond (fs, 10^{-15} s) timescale². Correspondingly, barrier crossing for a reaction in solution can happen on a timescale faster than most molecular motions occurring in the solvent. Although the 'continuum-solvent' picture provides a very appealing route to a description of solution-phase chemical dynamics, the equilibrium-based concept of a free-energy driving force becomes suspect; the gradient of a free-energy surface provides only the average force acting on the reactants. One must also ask how the instantaneous fluctuations in the force, which result from differences in solvent configuration, influence the outcome of an encounter between reactants. More generally, it is necessary to examine how the reactant and solvent dynamics are coupled. Considering the motion of a reacting molecule in solution, one must question to what extent such apparently microscopic characteristics of a dynamic reacting system can be productively viewed using macroscopic fluid concepts, and the potential importance of solvent molecularity becomes evident. Molecular solvent effects are far less well understood than are the potential influences of bulk properties. It is clear that the delineation of their role is a critical issue.

To fully address this issue, a detailed understanding is required of both the molecular motions that occur during the course of reaction, and the energy flow within the reacting molecule and between the reacting solute and the solvent. A general theoretical framework has been developed within which one can analytically describe the essential, but generic, roles of a solvent medium in determining condensed-phase reaction dynamics on an ultra-

short timescale³⁻⁵. Although a description of these developments is beyond the scope of this Review, these include elucidation of the critical role played by the solvent dynamics in determining reactant spatial diffusion during a reactive event and of the potentially rate-limiting role of solute-solvent energy transfer⁶⁻⁹. However, the area is not yet sufficiently developed that one can apply this framework directly and in general to the elucidation of the detailed molecular dynamics behind specific experimentally probed chemical systems.

Because of their ubiquitous appearance in all areas of chemistry and biochemistry, a class of reactions of special interest are those involving charge migration. Such reactions include not only electron and proton transfer, but also all of those in which the polarity of the reactants and products change significantly during the course of chemical change. These reactions commonly occur in polar liquids, and the corresponding strong interaction between polar solute and solvent induces an intimate coupling between them. Advances in laser technology are now enabling the direct experimental measurement of the dynamical behaviour of such chemical systems, allowing direct tests of theoretical models. At the same time, significant advances have been made in the ability to examine realistic theoretical models for these processes. Increasing computational power coupled with the parallel development of new computational algorithms allow for both classical and quantum-mechanical dynamical simulations of chemical phenomena on timescales accessible by these new experimental techniques. These efforts have resulted in an increasing synergism between computer simulation and laboratory experiment that is rapidly enhancing the molecular-level understanding of this important class of reactions. This review focuses on the new insights that are resulting from this interplay. Of course, simple collisional effects can also influence the course of the solution phenomena discussed. We do not discuss these aspects here as reviews on this subject can be found elsewhere¹⁰.

Experimental and theoretical methods

The invention of femtosecond laser spectroscopy enables the real-time study of dynamical processes in liquids. The general approach is to induce a change in the chemical solute using a short pulse of light, and then to interrogate the medium with a second pulse of light. Information carried in the time-dependence of the intensity, polarization or phase of the second laser pulse with respect to the photoinitiating light pulse provides information on the dynamical evolution of the reacting system. Various linear and nonlinear spectroscopies (which can involve more than two light pulses) have been developed to probe dynamical properties of chemical systems¹¹.

Despite this rapid growth in technological capabilities, there remain many limitations imposed on the chemical systems which

can be studied by ultrafast laser spectroscopy. Increased time resolution must result in decreased spectral resolution. To date, the shortest visible light pulses are of the order of 10 fs, corresponding to a spectral width of $\sim 1,000\text{ cm}^{-1}$. Whereas in principle it should be possible to carry out pump-probe experiments on a timescale that is pulse-width limited, most techniques used to study chemical systems have characteristic instrument responses of 50 fs or longer. The methods for extracting chemical rates from time-resolved data have been discussed in detail elsewhere¹².

Further, because laser techniques are ideal for initiating photochemical events and the wavelength range over which pulses can be generated is restricted, most experiments reported to date examine reactive processes of medium-sized molecules on excited-electronic-state surfaces. It is often the case that few details are known about the excited-state energy surfaces that govern the resulting chemical dynamics. Thus interpretation of experimental results is sometimes more difficult.

From a theoretical perspective, the focus lies more on the solvent. It can be described in simplest form in terms of macroscopic characteristics (viscosity and dielectric response) corresponding to a continuum description. Continuum-dielectric models for solvent effects on electrostatic free energy have a long history, and continue to play a major role in a number of areas¹³. In the simplest dynamical generalization, a reaction could be viewed as obeying diffusive dynamics subject to the mean force corresponding to the gradient of the free energy, but as pointed out above, such an assumed description is not valid *a priori*. The failure of the equilibrium mean force to satisfactorily describe the dynamics is sometimes referred to as "non-equilibrium solvation"⁵. Because continuum approaches explicitly lack molecularly, one must expect that an accurate description of solvent-induced local electric fields and rapid solvent configurational changes on the timescale of the individual molecular inertial motions (those occurring on short enough timescales that trajectories do not deviate significantly from force free motion, $\sim 10^{-14}$ s) are difficult within this framework. It is precisely in the case of ultrafast dynamics where these issues are most important.

One must then turn to fully molecular descriptions of both the solvent and solute, describing the solvent by explicit inclusion of a large number of such molecules. The intermolecular forces can then be specified by model potential functions, typically consisting of a sum of interatomic terms of the Lennard-Jones and Coulomb type which are derived based on a combination of quantum-chemical calculation and empirical results. Intramolecular (vibrational) degrees of freedom can also be included. In an area of active current development, an explicit quantum-chemical description of the solute electronic structure is employed¹⁴⁻¹⁶. Although aspects of the quantum behaviour of the solvent medium can be addressed¹⁷⁻¹⁹, it is a key feature of the systems of interest that quantum-mechanical issues reside predominantly in the electronic structural aspects of the solute, and recent developments emphasize a mixed quantum-solute/classical-solvent description of solutions.

Although analytical approximations can be implemented productively for fully molecular models²⁰, the dominant approach for elucidating molecular structure and dynamics resulting from solvent-solute interactions is the method of computer simulation²¹. In such approaches, the prescription of a potential-energy function leads explicitly to expressions for the energy and the forces prescribing the statistical configurational probabilities or the nuclear motions of both solute and solvent. In purely classical descriptions, the implementation of newtonian dynamics is routine. In the most recent approaches^{15,19}, both dynamics governed by a single (ground or excited) solute electronic-energy surface, and dynamics including electronic state-to-state transitions, have become accessible. Further recent developments have provided the tools for effective calculation of free energies²² and the direct evaluation of rate constants²³. It is important to emphasize that the useful implementation of simulation as an

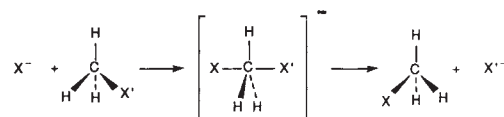
approach to these complex problems is attributable not only to the vast increase in accessible computational power but also to considerable theoretical work which underlies the new algorithms now permitting direct attack on these issues.

Chemical dynamics of polar systems

In this section, we emphasize (through example) that the solvent influences both the shape of the effective potential-energy surface for reaction and the motion on that surface, and thereby can act in a controlling manner in determining condensed-phase chemistry. These unifying concepts will be illustrated through the specific examples of nucleophilic substitution, solvation dynamics, electron and proton transfer, and electronic relaxation dynamics.

Nucleophilic substitution. We consider first the bimolecular nucleophilic substitution reaction, S_N2 , because it is a familiar and a well studied example that illustrates the key elements comprising the influence of a polar solvent on both the relevant reaction potential surface and on the dynamics of a true reactive event. Although time-resolved laser spectroscopy has not been used to examine the dynamics of the S_N2 reaction, the comparison of analytical theory with computer simulation provides important insights into the nature of the charge-transfer event. Most importantly, the elements identified here will be seen to underlie our understanding of the chemical processes discussed subsequently.

In a 'textbook' S_N2 reaction, and electron-rich reactant attacks the electron-deficient nuclear centre of a multiply substituted atom, leading to the replacement of one of the original substituents by the attacking group. In the prototype symmetric exchange reaction involving a halide ion (X^-) and a methyl halide molecule (CH_3X'), the halide attack on the tetrahedrally bonded carbon leads to an intermediate—or transition-state—structure, indicated in brackets, involving a five-coordinate carbon with the excess negative charge distributed over the complex:



The fact that the reactants, products and intermediate are charged leads to several strong effects of a polar solvent on the reaction. The electronic distribution in the reactant system is basically determined by the reactant nuclear positions, and is not strongly influenced by the distribution of surrounding solvent. (This is in contrast to the conclusions derived from studies of electron-transfer reactions, discussed later.) But because the charge is more localized in the reactants and products than in the transition state, the latter is relatively less well stabilized by interaction with the polar solvent. Hence the free-energy barrier to be overcome in solution is nearly twice as large in a polar solvent than in the gas phase.

Detailed treatments using molecular models of the type described above have examined the free-energy surface in solution for the prototypical symmetric exchange reaction of Cl^- with CH_3Cl , and reproduced the experimentally inferred solvent influence on the activation (free) energy, namely an increase by about 40 kJ mol^{-1} (refs 24, 25). Although the dominant effect of the solvent on the rate of reaction is due to this large change in free-energy barrier, solvent fluctuations play a significant role. The gas-phase potential is very steep near the transition state, and the transfer of charge from attacking to leaving group occurs very rapidly with only small geometric changes through the transition state (carrying the reactants from left to right in the above scheme). Simulations show²⁶ that, despite the relatively high mass of the reactant atoms, the solvent is nearly static on the timescale of this charge transfer (~ 20 fs). As a result, the final (reactive or non-reactive) fate of a reactant encounter is deter-

mined by a combination of the force intrinsic to the reaction plus the instantaneous (not mean) solvent force acting on the solute at the transition-state geometry^{5,26}. The latter can provide a sufficient restoring force opposing the charge migration so that reactant nuclear trajectories in the process of barrier crossing actually reverse direction about half of the time, never reaching the product state. The role of solvent fluctuations here is to change the rate by only about a factor of two. However, the effects of solvent exhibited for the S_N2 reaction form the foundation for understanding more complex chemical systems, as discussed below.

Solvation dynamics. Here solvation dynamics refers to studies that focus on the direct examination of the molecular relaxation dynamics of liquids that are subjected to a time-dependent perturbation from equilibrium. The essential relevance of these results lies in the fact that the dynamics of solvation of a varying charge distribution can influence the reaction rate for processes involving charge motion; this will be exemplified in subsequent sections.

We focus on time-dependent Stokes shift (TDSS) measurements of a dissolved solute molecule²⁷ where the interplay between experiment and simulation has had a major impact. TDSS studies involve monitoring the evolution of the emission spectrum of a dye molecule that is electronically excited by an ultrashort light pulse. Formation of the excited state occurs more rapidly than nuclear rearrangement of the solvent molecules. As a result, the excited-state molecule finds itself initially in a solvent configuration characteristic of the ground state. With increasing time, the solvent restructures in response to the charge distribution of the excited state of the probe molecule. This results in a time-dependent energetic stabilization of the excited molecule (and potentially a destabilization of the ground state), reflected in a corresponding dynamic redshift in the emission spectrum. The evolution of the energy of the system is quantified by a time-dependent function $C(t)$. Explicitly, if excitation is taken as occurring at time $t=0$, and denoting the average ground to excited state energy gap as $\Delta E(t)$ and the fully relaxed gap as $\Delta E(\infty)$, then one can define $C(t) = (\Delta E(t) - \Delta E(\infty)) / (\Delta E(t=0) - \Delta E(\infty))$, which yields a normalized time dependence for the relaxation process. Experimentally, $C(t)$ is evaluated from the evolution of a characteristic frequency (for example, the maximum or mean) of the solute emission spectrum; in computer simulations, $C(t)$ is typically calculated from the evolution of the interaction energy of a model for the ground- and excited-state solute, with the evolving electric potential deriving from the solvent.

One of the central assumptions in interpreting and generalizing TDSS results is that of a linear response²⁸. This embodies the idea that for a system close to equilibrium, one cannot distinguish between spontaneous fluctuations and deviations from equilibrium that are externally prepared. If assumed for the TDSS experiments, the measured (or calculated) solvation response can be directly related to equilibrium fluctuations in the liquid. Because it is equilibrium fluctuations that activate and deactivate solution-phase reactions, there is great interest in elucidating the timescales of these motions. Reported molecular-dynamics simulations^{29,34} indicate that the assumption of the linear response is reasonable for a number of solvent/solute systems, particularly for an aqueous solvent: clearly however, there must be limits to the applicability of this assumption. For example, some calculations³⁵ showed that linear response was unable to describe the solvation response of methanol to an instantaneously created model solute dipole. In this case, the early solvation dynamics are dominated by the local motions of OH bonds. With increasing time, other relaxation processes became important, most notably those taking place farther away from the solute molecule. But other recent calculations^{32,34} indicated that linear response can hold for solvation in methanol. It seems likely that the limitations on the accuracy of linear response for a given solvent depend on the size and charge

TABLE 1 Solvation times obtained from TDSS measurements

Solvent	Probe molecule	τ_s (ps)	$\tau_s/\tau_{\text{cont}}$
Acetonitrile	C102	0.9	4.0
	C311	0.7	3.0
	C152	0.52	2.5
Ethanenitrile	C102	1.5	4.2
	C311	1.1	3.1
	C152	0.85	2.4
Propionitrile	C102	1.5	2.8
	C311	1.6	2.9
Butyronitrile	C102	3.6	4.8
Cyanobenzene	C152	5.0	0.9
Methylacetate	C102	1.5	1.1
	C311	1.8	1.4
Ethylacetate	C102	2.7	1.3
	C311	2.3	1.1
Propylacetate	C102	4.0	1.5
Butylacetate	C102	6.6	1.8
	C152	0.68	2.0
Acetone	C153	1.03	3.0
	C152	1.2	0.7
Dimethylsulphoxide	C153	1.4	0.9
	C152	2.4	0.97
Propylene carbonate	C153	3.4	1.4

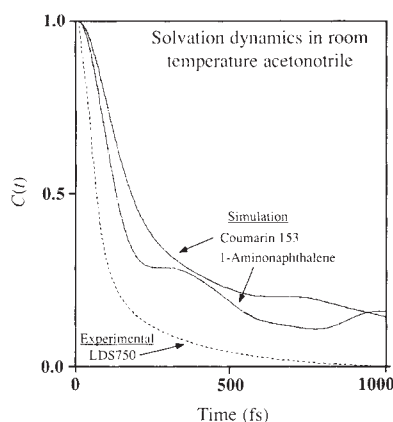
τ_s is the solvation time obtained from TDSS measurements; τ_{cont} is the predicted dielectric continuum relaxation time. Data taken from Table II of ref. 27; probe molecules are coumarin (C) dyes.

density of the model solute^{32,33}, and further studies are necessary to establish the degree of generality.

TDSS data have been recorded for many probe molecules in numerous solvents and at various temperatures²⁷. Table 1 gives the solvation times, τ_s , (given by the integral of the function $C(t)$ observed for various coumarin dye molecules in polar solvents, and also gives the ratio of τ_s to the relaxation time (τ_{cont}) predicted by dielectric continuum theory. In general, these studies reveal that for diverse cases the solvation time is similar to that predicted by dielectric continuum models. But the time dependence of the solvent response cannot be accounted for by simple bulk dielectric models: non-exponential solvent relaxation is observed for liquids that are adequately described by a (single timescale) Debye dielectric response. Coupled with the increasing recognition of the importance of solvent dynamical effects on reaction rates, the underlying physics behind this observed non-exponential time dependence of the solvation response needs to be understood. This has spurred several molecular dynamics simulations^{29,33,35,36} and a variety of molecularly based theories³⁷. These studies highlight the special importance of motions in the first solvent shell of the solute.

Simulation has played a critical role in motivating experimental studies accessing increasingly shorter time regimes. Data from simulations were the first to reveal large amplitude (>50%), ultrafast (tens of femtoseconds), initial components to the TDSS response^{29,30,35}. These results were interpreted in terms of librational motion in the case of water (a result seen in much earlier work³⁸) and the force-free (inertial) component of motion in a variety of polar aprotic solvents. But before such ultrafast behaviour could be considered important in solution-phase chemical reactivity, experimental verification was required. In 1991, Fleming and co-workers reported the experimental observation of an ultrafast relaxation component in the TDSS of the dye LDS-750 in acetonitrile³⁹. That component accounted for ~80% of the amplitude, and could be well described by a decay time of 70 fs. The slower tail, accounting for the remaining 20% of the relaxation dynamics, was exponential with a time constant

FIG. 1 The experimentally determined solvent correlation function, $C(t)$, for acetonitrile compared with that calculated from molecular-dynamics simulations. The agreement between calculation and experiment is nearly quantitative. Detailed analysis of the simulations reveals the molecular motions that are responsible for the solvent response.



of 200 fs. The experimental and simulated $C(t)$ functions (using models of the dye molecule coumarin 153 and of 1-aminonaphthalene as probes) for room-temperature acetonitrile are shown in Fig. 1. The agreement between experiment and simulation is striking. Analysis of the computer simulation data³⁰ revealed that the majority of the relaxation ($\sim 70\%$) is due to the small number of solvent molecules in the first solvent shell around the solute. The rapid fluctuations in the electric potential are correlated with rotational reorientation of the solvent molecules (~ 10 – 20°), not translational fluctuations in the solute–solvent distances. The simulations³⁰ also suggest that a major contribution to the relaxation comes from independent single-molecule motion. Very recently, corresponding theoretical and measured results have been reported⁸⁰ for aqueous C343, and these manifest very close agreement between the experimental $C(t)$ and the simulated linear response function.

Although the quantitative usefulness of the TDSS data in reaction rates remains to be further explored, these results provide evidence that the shortest-time solvent motions, and the immediate-neighbour molecules, have a dominant effect³⁶. These results demonstrate that future theoretical models of liquid-state dynamics must be based on molecular descriptions.

Electron and proton transfer. Electron-transfer reactions play a central role in chemistry and biochemistry. In particular, electron transfer forms the foundation of the field of electrochemistry, underlies the mechanisms of many organic processes, and is crucial to the primary steps of essential biological processes such as photosynthesis. As a result, there has been and continues to be a considerable theoretical and experimental effort aimed at understanding the role solvents play in determining reaction rates and efficiencies. A complete review of electron-transfer reactions is beyond the scope of this report; many excellent reviews and monographs have appeared^{40–44}.

The formation of a transition state in an electron-transfer reaction involves changes in the nuclear configuration of the reactants (inner-sphere reorganization) and/or the surrounding solvation shell(s) (outer-sphere reorganization). Limiting the discussion to the case where the outer-sphere contribution is large, it is fluctuations in the solvent that brings about occasional equalization of the energies of the reactant and product states. A schematic free-energy diagram is shown in Fig. 2. In the outer-sphere case, the reaction coordinate in the Figure corresponds to a displacement of the solvent, such as might be represented in the simplest case by a measure of the solvent polarization fluctuation. The free energy increases as the solvent deviates from the most probable configuration. The two displaced, intersecting, curves represent the free energy associated with each of the reference placements of the electron, on either acceptor or donor. At the instant in time when the solvent fluctuation brings the two electronic states into energetic coincidence, the electron can transfer without energetic penalty. This coupling of the

motion of the electron to the solvent nuclear degrees of freedom is in accord with the usual Born–Oppenheimer approximation. Insight into the timescales for these motions can be gained from TDSS studies (mentioned above). The rate at this point in time depends on the degree of electronic coupling between these reactant and product states. If the coupling is weak (the non-adiabatic limit), the probability for electron transfer is low. Here the electronic coupling, rather than solvent dynamics, will control the reaction rate. In the limit of strong coupling (the adiabatic limit), the reaction can be viewed as proceeding on a single electronic surface. Here the electron perfectly follows the nuclear motions of the solvent. The magnitude of coupling necessary for the reaction to be considered purely adiabatic is somewhat model-dependent. But it is generally accepted that for an adiabatic reaction the electron-transfer rate is directly proportional to the frequency of solvent fluctuations⁴⁵. Correspondingly, the solvent polarization can trap the electronic distribution in the vicinity of the top of the barrier, resulting in a reaction trajectory that involves multiple crossings through the transition state^{40,41}.

The applicability of these general theoretical concepts for understanding electron-transfer reactions have been validated by recent computer simulations^{46–49}. In particular, Chandler and co-workers calculated the relevant free-energy surfaces for the prototypical ferric–ferrous electron-transfer reaction in water⁴⁹. The calculated activation energy, 20 kcal mol^{-1} , was in good agreement with the estimated experimental value of 15 – 20 kcal mol^{-1} . More importantly, the data were quantitatively consistent with the Marcus picture⁴³ of parabolic free-energy surfaces (see Fig. 2) for electron-transfer processes.

Many recent experiments focus on the rate constants of excited-state intramolecular charge-transfer reactions⁵⁰. The molecules studied contain donor and acceptor groups which are covalently linked, enabling the study of electron-transfer kinetics in the absence of reactant diffusion. The reaction is initiated by the absorption of an ultrafast laser pulse, and therefore occurs on an excited-state potential. Despite the experimental advantages of this approach, an unfortunate result, noted earlier, is that the details of the reaction potential-energy surface are generally not available. In some cases it is not clear whether the reaction takes place on a single excited-state electronic surface, or if geometrical changes in the molecules associated with the charge-transfer process result in coupling two (or more) excited-state surfaces. Clearly there is the opportunity for considerable additional work on such systems.

Several studies report semiquantitative agreement between average solvation times determined from TDSS studies and the inverse of intramolecular charge-transfer rates, in cases with

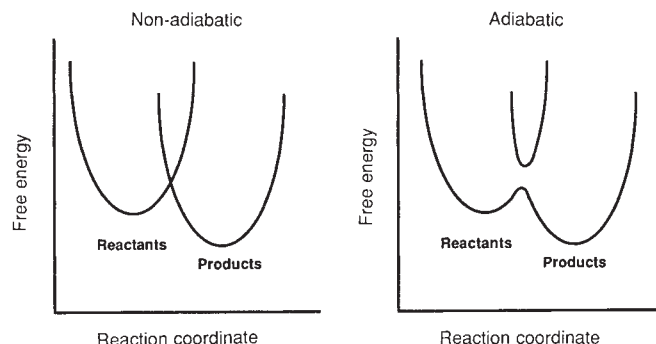


FIG. 2 Schematic reactant/product surfaces for non-adiabatic (left) and adiabatic (right) electron-transfer reactions. In the case considered in the text, the reaction coordinate involves changes in the solvent. In the non-adiabatic limit, the reactant and product surfaces do not mix and the reaction dynamics are not controlled by solvent motion. In the adiabatic limit, the potential-energy surfaces mix and the reaction trajectory effectively takes place on a single surface. In this limit, motion along the reaction coordinate is strongly coupled to fluctuations in the solvent.

reaction barriers less than kT (ref. 51). In particular, Barbara and co-workers reported quantitative agreements between inverse solvation times and the electron-transfer rates for bianthryl in a variety of polar solvents⁵¹. However, reaction rates that are faster than inverse solvation times have been reported for several molecular systems in a variety of solvents^{50,52}. For example, studies on the intramolecular charge transfer molecule DMAPS (bis-(*N,N*-dimethyl-aminophenyl)-sulphone) in alcohols revealed non-exponential kinetics, where the initial reaction rate was up to an order of magnitude faster than that predicted from average solvent times derived from TDSS studies. Two different explanations have been invoked to account for these observations. First, the reaction rate may be dominated by vibrational (or inner-sphere) dynamics. Second, the ultrafast component of solvent dynamics, not the average times revealed by the TDSS experiments, could dominate the results. The former issue has been discussed in detail by Sumi, Nadler and Marcus^{53,54}. The latter focus is the basis of time-dependent solvent models such as those proposed by Hynes⁵⁵. At present, it is difficult to distinguish experimentally between the importance of vibrational fluctuations and fast solvent dynamics in determining charge-transfer reaction rates that are faster than inverse solvation times. It is likely that both of these effects occur on the same timescale, and further experimental and theoretical work is necessary before their relative roles are quantified.

Many considerations similar to those addressed in the context of electron transfer also arise in the case of proton transfer^{23,56–59}. With respect to the role of solvent dynamics, proton transfer is expected to represent a difficult intermediate case between electron transfer and the heavy atom dynamics characteristic of the S_N2 reaction. Theoretical studies addressing the potential impact of the dynamics of the solvent on the rate of proton transfer for model potential surfaces have been reported^{23,56,60}, and suggest that the full range of solvent dynamical effects present in the charge-transfer cases already discussed are potentially present. From an experimental perspective, although some indication of solvent dynamical effects has been suggested⁶¹, clear experimental confirmation of such behaviour has not been obtained. In certain examples⁵⁸, the rate remains thermally activated even in strongly interacting solvents, whereas for others⁵⁹, the choice of solvent has a significant effect on the free energetics, but the reaction appears to take place faster than any characteristic solvent time. Part of the reason that experiments have not clearly shown dynamical solvent effects may reside in the focus on intramolecular proton-transfer reactions, where electronic redistribution can reduce the net amount of charge transfer and thus weaken the coupling to the solvent⁵⁶. Nevertheless, to the extent that intermolecular cases have been studied so far, an analysis based on free energies alone appears capable of correlating the available data⁶². Simulations pertinent to the molecules that have been addressed by ultrafast experiments are needed and are currently feasible, so that informative results are expected in the near future.

Electronic relaxation dynamics. Here we focus on the details of the relaxation that occurs in solution following an excitation of the electronic state of a solute, illustrating elements of both of the previous subsections. This relaxation includes both configurational rearrangement of the solvent, and return of the excited electronic-state of the solute to an equilibrium state. These two processes are intimately coupled, as the solvent dynamics can drive the electronic relaxation in the same way as discussed above for electron transfer. For the most part, we emphasize results related to the hydrated electron⁶³. Although the hydrated electron is a ubiquitous species in radiation chemistry and photochemistry, our interest here lies in the fact that it serves as a model for electronic-excited-state processes—this species has been the subject of intense recent experimental^{64–67} and theoretical^{15,19,68–71} attention, and it is now understood at a level of detail unavailable for any other system. We also briefly discuss related experimental results^{64,65,72,73} and calculations⁷⁴ on

the relaxation of excited-state ions in solution, which manifest some similar principles.

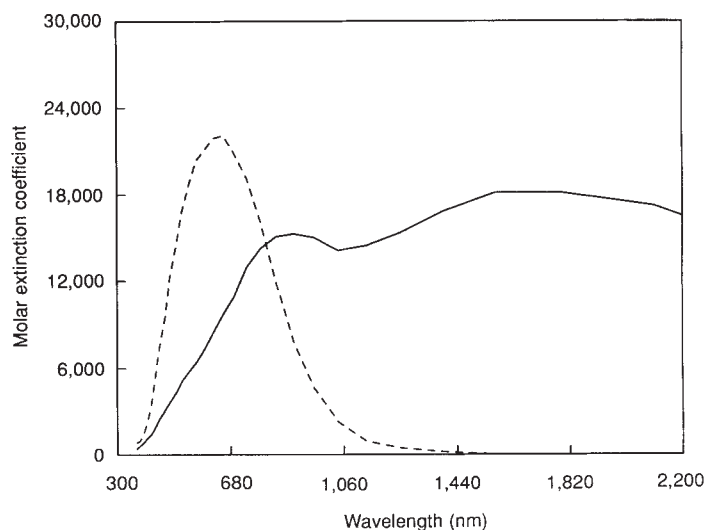
A collection of studies^{64,65} have focused on the sub-picosecond^{75,76} processes leading to an equilibrium ground-state hydrated electron, following photodetachment of an energetic electron from a water molecule in pure water. It was assumed until very recently that the loss of excess electronic energy was so rapid that observed dynamics would correspond primarily to the dynamics of solvation around a localized (ground-state) charge centre⁷⁵, analogous to the TDSS case discussed earlier, a hypothesis which may be tested experimentally. In particular, continuum-solvent descriptions⁷⁷ and molecular simulation⁶⁸ predict that the solvation of a localized electronic ground state should be signified by a single optical absorption band which manifests a continuous blue shift. In contradiction, femtosecond time-resolved experiments^{64,65} revealed transient spectra involving at least two distinct bands, demonstrating that the electron solvation process involves at least two distinct electronic states. A proposal⁶⁸ that the ultrafast component of solvation could allow solvation dynamics to compete with electronic energy loss, leading to a solvated excited state as a well defined intermediate, was confirmed by later simulations of the process^{15,70} and corresponding spectral analysis⁷¹.

The existence of detailed theoretical analyses of the excited states of the hydrated electron, and the existence of improved laser technology, have now permitted an even more direct probe of the details of solvation and electronic dynamics involving the ground state and first excited state of this system. In summary⁶⁸, the hydrated electron is characterized by a nearly spherical (*s*-like) ground-state electron distribution, with solvent fluctuations leading to a range of instantaneous energies, and a triplet of localized (*p*-like) excited states that are non-degenerate due to the instantaneous asymmetry of the surrounding solvent. The higher-lying states form a delocalized continuum. The corresponding ground-state absorption spectrum is dominated by transitions to the *p*-like states. Calculations⁷⁰ of the ground-state spectrum (in reasonable accord with experiment) and the spectrum of an electron equilibrated in the first excited (*p*-like) state are shown in Fig. 3. As can be seen in the Figure, the excited-state spectrum is considerably redshifted, in accord with the implications of earlier analysis of the intermediate state evident in experimental photoionization data⁶⁵.

If the set of ground-state hydrated electrons resonant with a narrow-bandwidth pump laser are promoted into the excited state^{66,67}, there will be several expected contributions to the observed transient spectra. An (initially non-uniform) bleach of the ground-state absorption will appear due to the loss of a subpopulation resonant with the pump; bleach homogeneity will appear on the timescale of solvent fluctuations. An excited-state absorption contribution will also contribute, subject to transient evolution as the solvent relaxes around the new electronic distribution. Further, the excited state has a characteristic rate for electronic energy loss to the solvent, and the resulting ground state is then subject to solvent relaxation, returning to its original equilibrium state. In recent experiments^{66,67}, electrons were detached from iodide ions by an initial ultraviolet pulse (a pure water sample yields the same results), and after a delay of the order of a nanosecond, the equilibrium hydrated electrons are subject to this pump-probe experiment, using a pump wavelength of 780 nm, roughly at the crossing (isosbestic) point of the ground- and excited-state spectra shown in Fig. 3.

Based on Fig. 3, in the absence of solvent dynamical contributions, one would expect the excitation to induce a transient absorption to the red and a transient bleach to the blue of the pump wavelength, both relaxing back to equilibrium on the timescale of the excited-state lifetime. What is seen in experiment^{66,67} and in simulation⁶⁹ is illustrated in Fig. 4 for three representative wavelengths. Whereas the results at the red and blue extremes are as expected, the intermediate wavelengths show a transient bleach followed by a transient absorption.

FIG. 3 Calculated absorption spectra for a model of the hydrated electron for the equilibrium ground state (dashed line) and for the equilibrated first excited state (solid line). The result implies that transient excitation to the first excited state would result in a transient bleach to the blue, and a transient absorption to the red, of ~ 800 nm. (The spectra calculated for the models discussed here (see ref. 67) are subject to a blue shift compared to the experiment. To give the wavelengths for comparison to experiment, all results reported employ a single constant energetic shift (0.62 eV) of the directly calculated energies. The same correction is applied to the calculated spectra in Fig. 4.)



Simulation⁶⁹ reveals that this behaviour reflects solvent dynamics, primarily excited-state solvent relaxation and evolution of the ground-state bleach. We note that the experimental data can also be fitted by phenomenological spectral modelling that includes solvent dynamics only through solvent relaxation around the final ground state⁶⁷, although a significant contribution from this effect is not evident in the simulation results considered here.

A feature not previously discussed here is of special importance: the excited-state lifetime is also strongly dependent on solvent dynamics. In addition to the coupling of ground and excited electronic states by the solvent motion¹⁵, which we do not discuss here, it is well known that the transition rate between states decreases with the energy gap between them⁷⁸. For the present case, the initial energy gap on excitation is about 1.6 eV (780 nm), yielding a relatively slow initial transition rate, with the rate increasing as this gap narrows to only about 0.4 eV as the solvent relaxes around the excited state. The dynamics of this enormous Stokes shift is displayed in Fig. 5 through a direct calculation of the decay function $C(t)$ discussed above. About half of the gap reduction is seen to occur on the inertial time scale of about 25 fs and the remainder on a time scale of roughly

250 fs. These results are strikingly similar to recent TDSS measurements for molecular solvation in water⁸⁰. Also shown in Fig. 5 is the prediction based on the linear-response assumption that this nonequilibrium relaxation follows the decay of fluctuations in the gap for the equilibrium ground state. For this case, the linear-response hypothesis is seen explicitly to be a very good one for describing the Stokes shift.

Finally, we consider briefly a more chemical example. The dynamics of photodetachment of an electron from halide ions in water^{64,65,72-74} also indicates the importance of an ultrafast solvent-driven solute electronic relaxation process. The results of classic experiments in this area⁷⁹ showed that in solution, such halide ions exhibit a strong absorption in the ultraviolet regime associated with so-called charge-transfer-to-solvent (CTTS) states; these states depend on the presence of a polar solvent medium for their existence. Recent experiments from two groups have provided femtosecond time-resolved spectroscopy of photoexcited chloride^{64,65} and iodide^{72,73}. Experimental results for iodide indicate the resolution at very short times of a transient species, tentatively identified as the excited atomic CTTS state⁷². Recent simulation studies⁷⁴ indicate that a CTTS state lying above the principal band is directly excited in the experiment (8 eV excitation), and that an experimentally unresolved ultrafast (≤ 50 fs) relaxation to the lowest excited-state band occurs before further evolution. The most recent experimental results⁷³ are consistent with this view.

Complete relaxation of the solvent for the lowest excited state of the ion leaves a remaining gap to the ground state of several electron volts. Direct simulation of the further dynamics⁷⁴ has revealed a novel principal pathway by which ground-state recovery occurs; it is not by direct disposal of this large amount of energy into the solvent, but rather by the much more rapid

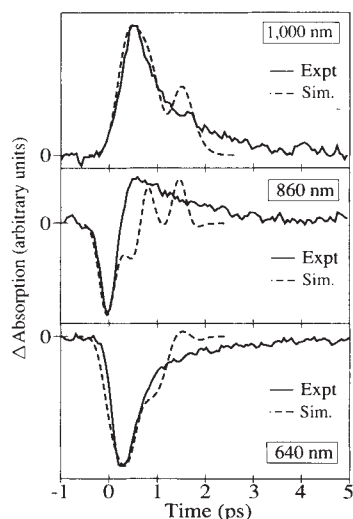


FIG. 4 Differential absorption spectra at indicated probe wavelengths for the hydrated electron after excitation by a 780-nm pump laser of the equilibrium ground state, producing a p -like excited state. Solid lines, experimentally determined⁶⁷; dashed lines, calculated from simulations⁶⁸. The data are normalized to the same maximum amplitude.

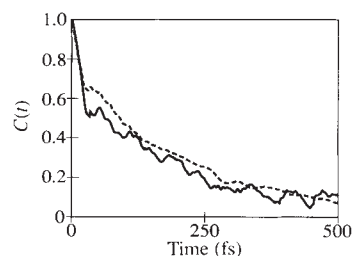


FIG. 5 Time dependence of the relaxation of the energy gap between the ground state and first excited state of the hydrated electron following initial excitation. Solid line, calculated (nonequilibrium) result; dashed line, linear-response prediction based on fluctuations in the ground state.

process of a solvent-driven electron detachment to form a vicinal ground-state hydrated electron, followed by geminate recombination with the parent atom (solvent-to-atom electron transfer) to reform the ground-state ion in only a few picoseconds. Such solvent-driven electronic dynamics is probably a widespread mechanism for rapid non-radiative relaxation which has not previously been identified.

Future prospects

The examples outlined above have shown that in several cases the detailed microscopic mechanisms of polar solution chemistry intimately involve the solvent in its description. The thermodynamic and time-dependent influence of the solvent is crucial to the rate of reactions, and the modern experimental and computational attack is allowing us to discern features at a molecular level and to develop generalizable principles governing these processes. As this evolution continues, one should anticipate growth on two fronts. First, although the reactions discussed here are important ones, they are also relatively simple. It is expected that the framework for understanding solution chemistry that is engendered by these studies will make easier the analysis of considerably more complex

reaction systems. An important class of examples includes photoinitiated bond formation and molecular photodissociation. Second, one can ask even more detailed questions about the role of solvent. For example, by direct analogy with the theory of molecular non-radiative relaxation, it is known that solvent motions stimulate electronic relaxation and that the energy is dissipated in the solvent. Such processes must compete with both potentially interesting, and, in other cases, undesirable photoinitiated reactions. But the specifics of such solvent dynamical coupling is essentially unexplored and the ability to interpret experimental data on such rates at a molecular level is only now beginning. With the combined growth of both computational and analytical theory and of experimental techniques, the elucidation of the detailed dynamics of both ground- and excited-state processes on a much wider class of solution systems will make rapid progress. □

Peter J. Rossky is at the Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, Texas 78712-1167, USA. John D. Simon is at the Department of Chemistry, University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093-0341, USA.

- Reichardt, C. *Solvents and Solvent Effects in Organic Chemistry* (VCH, Weinheim, 1988).
- Gruebele, M. & Zewail, A. H. *Physics Today* **43**, 24–33 (1990).
- Hänggi, P., Talkner, P. & Borkovec, M. *Rev. mod. Phys.* **62**, 251–341 (1990).
- Berne, B. J., Borkovec, M. & Straub, J. E. *J. phys. Chem.* **92**, 3711–3725 (1988).
- Hynes, J. T. A. *Rev. phys. Chem.* **36**, 573–597 (1985).
- Grote, R. F. & Hynes, J. T. *J. chem. Phys.* **73**, 2715–2732 (1980).
- Carmeli, B. & Nitzan, A. *Phys. Rev. Lett.* **49**, 423–426 (1982).
- Eads, D. D., Dismer, B. G. & Fleming, G. R. *J. chem. Phys.* **93**, 1136–1148 (1990).
- Frishman, A. M. & Pollak, E. J. *chem. Phys.* **98**, 9532–9543 (1993).
- Fleming, G. R., Courtney, S. G. & Balk, M. W. *J. Statist. Phys.* **42**, 83–104 (1986).
- Mourou, G. & Zewail, A. H. (eds) *Ultrafast Phenomena VIII* (Springer, New York, 1994).
- Simon, J. D. (ed.) *Ultrafast Dynamics of Chemical Systems* (Kluwer, Dordrecht, 1994).
- Sharp, K. & Honig, B. A. *Rev. Biophys. biophys. Chem.* **19**, 301–332 (1990).
- Remler, D. K. & Madden, P. A. *Molec. Phys.* **70**, 921–939 (1990).
- Webster, F. J., Schnitker, J., Friedrichs, M. S., Friesner, R. A. & Rossky, P. J. *Phys. Rev. Lett.* **66**, 3172–3175 (1991).
- Field, M. J., Bash, P. A. & Karplus, M. *J. comput. Chem.* **11**, 700–733 (1990).
- Del Buono, G. S., Rossky, P. J. & Schnitker, J. J. *chem. Phys.* **95**, 3728–3737 (1991).
- Bader, J. S., Kuharski, R. A. & Chandler, D. J. *chem. Phys.* **93**, 230–236 (1990).
- Neria, E., Nitzan, A., Barnett, R. N. & Landman, U. *Phys. Rev. Lett.* **67**, 1011–1014 (1991).
- Rossky, P. J. A. *Rev. phys. Chem.* **36**, 321–346 (1985).
- Whitnell, R. M. & Wilson, K. R. *Reviews in Computational Chemistry* (eds Lipkowitz, K. B. & Boyd, D. B.) 67–148 (VCH, New York, 1993).
- Beveridge, D. L. & DiCapua, F. M. A. *Rev. Biophys. biophys. Chem.* **18**, 431–492 (1989).
- Borgis, D., Tarjus, G. & Azzouz, H. *J. phys. Chem.* **96**, 3188–3191 (1992).
- Chandrasekhar, J., Smith, S. F. & Jorgensen, W. L. *J. Am. chem. Soc.* **107**, 154–163 (1985).
- Huston, S. E., Rossky, P. J. & Zichi, D. A. *J. Am. chem. Soc.* **111**, 5680–5687 (1989).
- Bergsma, J. P., Gertner, B. J., Wilson, K. R. & Hynes, J. T. *J. chem. Phys.* **86**, 1356–1376 (1987).
- Maroncelli, M. *J. Molec. Liquid* **57**, 1–37 (1993).
- Chandler, D. *Introduction to Modern Statistical Mechanics* (Oxford Univ. Press, New York, 1987).
- Maroncelli, M. & Fleming, G. R. *J. chem. Phys.* **89**, 5044–5069 (1988).
- Maroncelli, M. *J. chem. Phys.* **94**, 2084–2103 (1991).
- Carter, E. A. & Hynes, J. T. *J. chem. Phys.* **94**, 5961–5979 (1991).
- Ando, K. & Kato, S. *J. chem. Phys.* **95**, 5966–5982 (1991).
- Maroncelli, M. et al. in *Ultrafast Reaction Dynamics and Solvent Effects* (eds Gauduel, Y. & Rossky, P. J.) 310–333 (American Institute of Physics, New York, 1994).
- Zhu, J. & Cukier, R. I. *J. chem. Phys.* **98**, 5679–5693 (1993).
- Fonseca, T. & Ladanyi, B. M. *J. phys. Chem.* **95**, 2116–2119 (1991).
- Neria, E. & Nitzan, A. *J. chem. Phys.* **96**, 5433–5440 (1992).
- Bagchi, B. A. *Rev. phys. Chem.* **40**, 115–141 (1989).
- Rao, M. & Berne, B. J. *J. phys. Chem.* **85**, 1498–1505 (1981).
- Rosenthal, S. J., Xie, X., Du, M. & Fleming, G. R. *J. chem. Phys.* **95**, 4715–4718 (1991).
- Hynes, J. T. *J. statist. Phys.* **42**, 149–168 (1986).
- Chandler, D. *J. statist. Phys.* **42**, 49–67 (1986).
- Fox, M. A. (ed.) *Chem. Rev.* **92**(3), 365–490 (1992).
- Marcus, R. A. A. *Rev. phys. Chem.* **15**, 155–196 (1964).
- Sutin, N. *Prog. inorg. Chem.* **30**, 441–498 (1983).
- Weaver, M. J. *J. Chem. Rev.* **92**, 463–480 (1992).
- Hwang, J.-K. & Warshel, A. *J. Am. chem. Soc.* **109**, 715–720 (1987).
- Zichi, D., Cicciotti, G., Hynes, J. T. & Ferrario, M. J. *phys. Chem.* **93**, 6261–6265 (1989).
- Kuharski, R. A. et al. *J. chem. Phys.* **89**, 3248–3257 (1989).
- Bader, J. S. & Chandler, D. *Chem. Phys. Lett.* **157**, 501–504 (1989).
- Barbara, P. F., Walker, G. C. & Smith, T. P. *Science* **256**, 975–981 (1992).
- Kang, T. J., Jarzeba, W., Barbara, P. F. & Fonseca, T. *Chem. Phys.* **149**, 81–95 (1990).
- Simon, J. D. & Doolen, R. J. *Am. chem. Soc.* **114**, 4861–4870 (1992).
- Sumi, H. & Marcus, R. A. *J. chem. Phys.* **84**, 4894–4914 (1986).
- Nadler, W. & Marcus, R. A. *J. chem. Phys.* **86**, 3906–3924 (1987).
- Hynes, J. T. *J. phys. Chem.* **90**, 3701–3706 (1986).
- Timoneda, J. J. I. & Hynes, J. T. *J. phys. Chem.* **95**, 10431–10442 (1991).
- Scheiner, S. *Accs chem. Res.* **18**, 174–180 (1985).
- Barbara, P. F., Walsh, P. K. & Brus, L. E. *J. phys. Chem.* **93**, 29–34 (1989).
- Smith, T. P. et al. *J. phys. Chem.* **95**, 10465–10475 (1991).
- Borgis, D. & Hynes, J. T. *J. chem. Phys.* **94**, 3619–3628 (1991).
- Kim, Y. R., Yardley, J. T. & Hochstrasser, R. M. *Chem. Phys.* **136**, 311–319 (1989).
- Pines, E. & Fleming, G. R. *J. phys. Chem.* **95**, 10448–10457 (1991).
- Jortner, J. & Kestner, N. R. (eds) *Electron-Solvent and Anion-Solvent Interactions* (Elsevier, Amsterdam, 1976).
- Gauduel, Y., Pommeret, S., Migus, A., Yamada, N. & Antonetti, A. *J. opt. Soc. Am.* **B7**, 1528–1539 (1990).
- Lu, H., Long, F. H. & Eisenthal, K. B. *J. opt. Soc. Am.* **B7**, 1511–1520 (1990).
- Alfano, J. C., Walhout, P. W., Kimura, Y. & Barbara, P. F. *J. chem. Phys.* **98**, 5996–5998 (1993).
- Kimura, Y., Alfano, J. C., Walhout, P. W. & Barbara, P. F. *J. phys. Chem.* **98**, 3450–3458 (1994).
- Rossky, P. J. & Schnitker, J. J. *J. phys. Chem.* **92**, 4277–4285 (1988).
- Schwartz, B. J. & Rossky, P. J. *J. phys. Chem.* **98**, 4489–4492 (1994).
- Murphrey, T. H. & Rossky, P. J. *J. chem. Phys.* **99**, 515–522 (1993).
- Keszei, E., Nagy, S., Murphrey, T. H. & Rossky, P. J. *J. chem. Phys.* **99**, 2004–2011 (1993).
- Long, F. H., Lu, H., Shi, X. & Eisenthal, K. B. *Chem. Phys. Lett.* **169**, 165–171 (1990).
- Long, F. H., Shi, X., Lu, H. & Eisenthal, K. B. *J. phys. Chem.* (in the press).
- Shen, W.-S. & Rossky, P. J. *J. Chem. Phys. Lett.* **213**, 233–238 (1993).
- Kennedy-Wallace, G. A. & Jonah, C. D. *J. phys. Chem.* **86**, 2572–2586 (1982).
- Wiesenfeld, J. M. & Ippen, E. P. *Chem. Phys. Lett.* **73**, 47–50 (1980).
- Calef, D. F. & Wolynes, P. G. *J. chem. Phys.* **78**, 4145–4152 (1983).
- Freed, K. F. in *Radiationless Processes in Molecules and Condensed Matter*, (ed. Fong, F. K.) (Springer-Verlag, New York, 1976).
- Blandamer, M. J. & Fox, M. F. *Chem. Rev.* **70**, 59–93 (1970).
- Jimenez, R., Fleming, G. R., Kuman, P. V. & Maroncelli, M. *Nature* **369**, 471–473 (1994).

ACKNOWLEDGEMENTS. This work was supported by the US National Science Foundation (J.D.S., P.J.R.), the US National Institutes of Health (J.D.S., P.J.R.), the US Office of Naval Research (J.D.S., P.J.R.), the R. A. Welch Foundation (P.J.R.) and the Texas Advanced Research Program (P.J.R.). We thank P. Barbara for helpful discussions, and M. Maroncelli for providing Fig. 1 and for many helpful discussions.

Structure and mechanism of interleukin-1 β converting enzyme

Keith P. Wilson, Jo-Anne F. Black, John A. Thomson, Eunice E. Kim, James P. Griffith, Manuel A. Navia, Mark A. Murcko, Stephen P. Chambers, Robert A. Aldape, Scott A. Raybuck & David J. Livingston

Vertex Pharmaceuticals Incorporated, 40 Allston Street, Cambridge, Massachusetts 02139, USA

Interleukin-1 β converting enzyme (ICE) processes an inactive precursor to the proinflammatory cytokine, interleukin-1 β , and may regulate programmed cell death in neuronal cells. The high-resolution structure of human ICE in complex with an inhibitor has been determined by X-ray diffraction. The structure confirms the relationship between human ICE and cell-death proteins in other organisms. The active site spans both the 10 and 20K subunits, which associate to form a tetramer, suggesting a mechanism for ICE autoactivation.

HUMAN interleukin-1 β converting enzyme (ICE), a cysteine protease localized primarily in monocytes, processes the 33K inactive precursor of interleukin-1 β (pIL-1 β) to the 17K pro-inflammatory cytokine IL-1 β (refs 1, 2). ICE has a unique sequence preference among mammalian proteases, and requires substrates with aspartic acid adjacent and n-terminal to the scissile peptide bond (P₁ residue)^{3,4}. To date, pIL-1 β is the only identified physiological substrate. pIL-1 β lacks a conventional leader sequence and is not processed by a signal peptidase⁵, nor directed to the Golgi apparatus for secretion⁶. Instead, pIL-1 β is cleaved by ICE between Asp 116 and Ala 117 to produce the biologically active carboxy-terminal fragment found in human serum and synovial fluid^{3,4}. Processing by ICE is necessary for transport of mature IL-1 β through the cell membrane.

Active ICE is an oligomeric enzyme with subunits of relative molecular masses (M_r) of 20K and 10K, denoted here as p20 (residues 120–297) and p10 (317–404), respectively^{7,8}. The subunits are derived from a 45K pro-enzyme (p45) by cleavage at four Asp–Xaa bonds, Asp 103–Ser 104, Asp 119–Asn 120, Asp 297–Ser 298, and Asp 316–Ala 317 (ref. 7). Purified p45, and the p30 form which lacks the amino-terminal pro sequence, can autoprocess to the active form *in vitro* by manipulation of enzyme concentration and temperature⁷.

ICE or ICE homologues may regulate programmed cell death (apoptosis). ICE shares 28% sequence identity with CED-3, a *Caenorhabditis elegans* protein in the apoptosis pathway, and overexpression of murine ICE in *Rat-1* fibroblast cells causes cell death^{9,10}. Increased ICE activity has been reported in phagocytic cells stimulated with agents which promote apoptosis¹¹. The *crmA* protein, a viral serpin inhibitor of ICE¹², protects transfected ganglion neurons from apoptosis induced by nerve growth factor depletion¹³. These observations suggest that programmed cell death observed in degenerative neuronal diseases, such as Alzheimer's and Parkinson's diseases¹⁴, may be regulated by ICE or ICE homologues.

We have solved the structure of ICE at 2.6 Å resolution by X-ray diffraction (Table 1). Structural, kinetic and mutagenesis data suggest mechanisms for ICE catalysis and regulation.

Fold and topology of ICE

Each ICE heterodimer resembles a rectangular box with rough dimensions 45 Å × 35 Å × 25 Å. The p20 and p10 subunits are intimately associated, and the protein appears as a single domain (Fig. 1). The enzyme core is a six-stranded β -sheet with five parallel strands (1, residues 164–170; 2, 199–205; 3, 230–236; 4,

278–283; and 7, 327–331) and one anti-parallel strand (8, residues 388–393). The anti-parallel strand lies on one edge of the sheet, and is perpendicular and adjacent to the crystallographic 2-fold axis. Six α -helices (A, residues 139–149; B, 182–196; C, 207–226; D, 257–269; E, 348–363; and F, 366–377) lie roughly parallel to the β -strands and shield the β -sheet core from solvent. The last seven residues of p20 and the first seven of p10 protrude from this compact structure and form two anti-parallel β -strands (5, residues 291–297 and 6, 317–323) that interact extensively across the 2-fold axis of the crystal (Fig. 2).

The active site

We solved the structure of ICE bound to a tetrapeptide aldehyde inhibitor (acetyl-Tyr-Val-Ala-Asp-H)¹⁵ to suppress autoproteolysis at the protein concentrations used for crystallization. The inhibitor is covalently bound to the sulphur atom of Cys 285 (Fig. 4a). Other cysteine-modification agents are selective for Cys 285 (ref. 7). In concert with the site-directed mutagenesis experiments reported here (Fig. 3), Cys 285 is confirmed as the active site nucleophile. The His 237 imidazole side chain is adjacent to Cys 285 (Fig. 4a) and participates in a hydrogen-bond network that includes the oxyanion of the aldehyde inhibitor, suggesting a role for this residue in catalysis.

The structure of the inhibited complex reveals why both p20 and p10 subunits are required for ICE activity⁷. The tetrapeptide inhibitor binds in the S₁ to S₄ ICE subsites which would contact peptide substrate residues N-terminal to the scissile peptide bond. Side chains from p10 form the S₂ to S₄ subsites (Fig. 4b). The inhibitor makes main chain to main chain contacts with residues from p10 in an anti-parallel β -sheet arrangement (Fig. 4b). The side chains of p10 residues Val 338 to Pro 343 interact with the inhibitor, except for Ser 339, whose side chain is buried in the protein core. The γ -oxygen of Ser 339 is near the sulphur atom of Cys 285 (3.9 Å), and is stabilized by the side-chain oxygen atoms of two other buried Ser residues (332 and 333) and a water molecule. Both subunits contribute to the selective aspartic acid recognition by the S₁ subsite⁷, through direct charge-charge interactions with the side chains of Arg 179 (p20) and Arg 341 (p10). On the basis of hydrogen bonding distances and geometry, the more important interaction is with residue Arg 179 (Fig. 4a). The side chain NH of Gln 283 also hydrogen bonds to the P₁ carboxylate side chain of the inhibitor. The Ala side chain at P₂ and the Val at P₃ of the inhibitor are exposed to solvent (Fig. 4b), consistent with the broad tolerance of substitution in the P₂ position of substrates⁷. The Tyr side chain in P₄ contacts the

FIG. 1 Stereo ribbon⁴³ drawing of the p20 (blue)p10 (red) ICE heterodimer. Secondary structural elements are labelled (gold lettering) as defined in the text. The placement and orientation of secondary structural elements in ICE is similar to a portion of horse liver alcohol dehydrogenase and thioredoxin^{44,45}. A few key residues referred to in the text are labelled (green lettering) according to their position in the p45 amino-acid sequence of ICE^{7,8}. The locations of site-directed mutants discussed in this report are highlighted (green backbone) and have their side chain atoms displayed. The active site is at the top of the figure, roughly at the centre of the cluster of displayed side chains. The side chains of His 237 and the catalytic Cys 285 are coloured purple. From left to right across the top of the figure are His 248, Cys 244, His 237, Arg 179, Cys 285 and the serine triplet, Ser 332, Ser 333 and Ser 339 (closest to Cys 285). A disulphide bond between Cys 136 from p20 and Cys 362 from p10 is observed (shown at lower right). Cys 136 is conserved in the murine and rat ICE sequences, but Cys 362 is not. Because disulphide bonds in cytoplasmic proteins are rare, this covalent modification may be an artefact of purification, refolding or crystallization. Mutation of ICE



residue Cys 136 or Cys 362 to Ser maintains enzymatic activity, demonstrating that disulphide bond formation is not essential for enzymatic activity (Fig. 3).

side chains of His 342, Val 348 and Arg 383 (Fig. 4b), but it is not hydrogen bonded to the enzyme. Modelling studies indicate that the C-terminal region of substrate would contact p20 residues exclusively (in the S₁' to S₄ subsites).

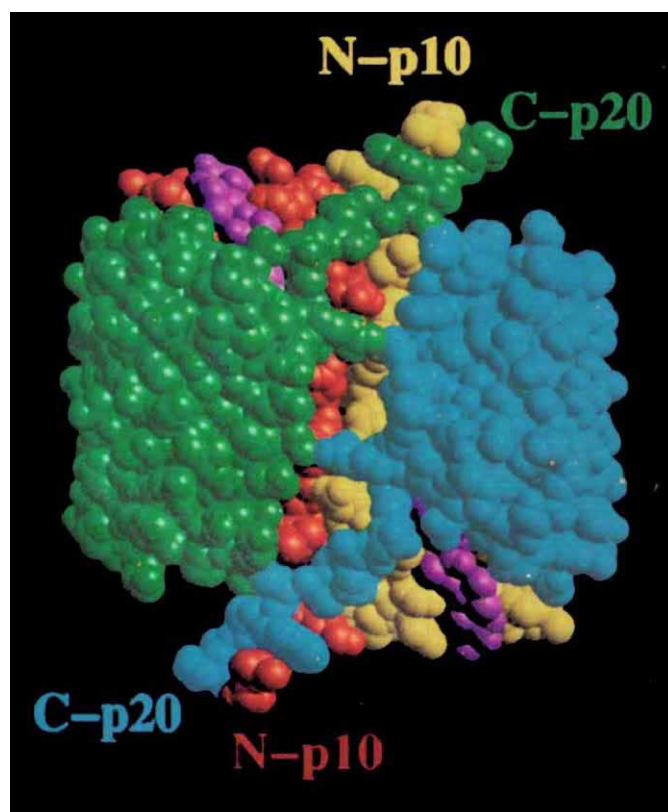


FIG. 2 CPK model of the (p20)₂/(p10)₂ tetramer of ICE. Two p20 subunits (green and blue) surround two adjacent p10 subunits (red and gold). The tetrapeptide aldehyde inhibitor is coloured purple. The crystallographic 2-fold axis is roughly perpendicular to the plane of drawing, and runs through the small hole at the centre of the interface between the two p10 subunits. Beta-strands 5 and 6 protrude from the top, and by 2-fold symmetry, from the bottom of the model. The N- and C-terminal end of each subunit is labelled. Dynamics simulations suggest that formation of a tetramer would stabilize the active site of ICE by reducing the mobility of the catalytic Cys 285 and other residues near the p20 C-terminus (D. Pearlman and M.A.M., unpublished results).

A second Cys/His pair (Cys 244–His 248) is found near the active site. The side chains of these two residues are in van der Waals contact, and these residues may contact the P₂' and P₃' side chains of substrates. Cys 244 also contacts His 237, the active site His. Together, the four residues are arranged in a side-by-side manner, Cys–His–Cys–His (Fig. 1). There is no evidence that these residues coordinate divalent metal ions, but thimerosal, a Hg compound, modifies Cys 244 in the presence of the aldehyde inhibitor and is a 15 μ M uncompetitive inhibitor of ICE activity.

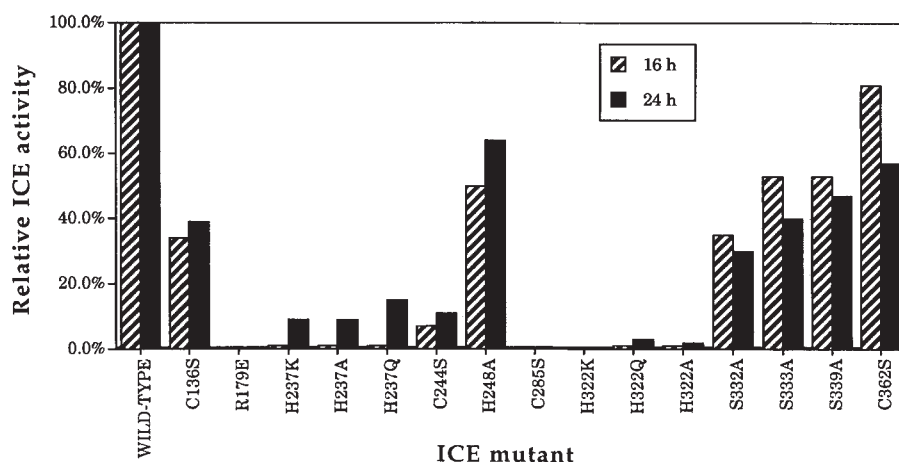
On the basis of an alignment of the ICE sequence Gly 287–Asp–Ser–Pro–Gly 291 to a consensus sequence in serine and viral cysteine proteases, it was suggested that Ser 289 might play a catalytic role in ICE⁷. The structure, however, shows that these residues form a turn near the C terminus of p20 and are not involved in catalysis. Mutation of Gly 287 to Ser eliminates the ability of murine ICE to induce apoptosis⁹. Both Gly residues of this pentapeptide have $\phi\psi$ angles that are excluded for amino acids with bulkier side chains. Gly 287 may also be involved in recognition of the P' end of substrate. Thus, substitution at Gly 287 is likely to disrupt the structure near the active site.

Mutational and kinetic analysis of the active site

Inhibitors are used routinely to assign proteases to mechanistic classes, but ICE is not easily categorized in this way. Cys modification reagents such as pCMB, iodoacetate⁷ and iodoacetamide¹⁶, and the His modification agent DEPC (diethyl pyrocarbonate), inactivate ICE, and substrate protects against inactivation by these compounds (data not shown). These results suggest a role for active site Cys and His residues. Yet ICE is insensitive to several classic cysteine protease inhibitors, such as E64 and antipain. ICE is also inactivated by TPCK (*N*-tosyl-L-phenylalanine chloromethyl ketone) and 3,4-dichloroisocoumarin, reagents considered to be selective for serine proteases. Taken together, these observations suggest differences in the binding sites of ICE and papain, the latter a well-studied cysteine protease of known structure with high sequence similarity in the active site region to the human cysteine proteases cathepsin B and H (ref. 17).

To explore further the roles of ICE active site residues in substrate binding and catalysis, we investigated the ability of a series of p30 ICE mutants to process pIL-1 β *in vivo* (Fig. 3). We transfected each ICE mutant into a COS cell line that constitutively expresses pIL-1 β , and measured the amount of mature IL-1 β secreted. Enzymatic activity in this assay requires that transcription, translation and protein folding of mutants are not arrested. We determined, therefore, the amount of mutant ICE

FIG. 3 Activity of ICE mutants in processing pIL-1 β in COS-1 cells relative to wild-type activity. Mutations were made in the 30K ICE-encoding cDNA cloned into a pcDNA3 mammalian expression vector (Invitrogen) by oligonucleotide-directed mutagenesis, and fully sequenced in the coding region. The COS-1 cell line had been previously transfected with pIL-1 β encoding cDNA which had integrated into the chromosome. pIL-1 β production from COS-1 cells was maintained by the addition of 0.5 mg ml⁻¹ G-418 sulphate to culture media. About 3×10^6 COS-1 cells in 100-mm tissue culture plates were transfected with 15 μ g of each plasmid. DNA was mixed with 200 μ l DEAE-Dextran, brought to 4 ml with phosphate-buffered saline, and added to the plates. Cells were incubated at 37 °C for 30 min. 8 ml of an 80 μ M chloroquine/serum-free DMEM solution was added and cells were incubated for 2.5 h. This solution was aspirated and cells were treated for 2 min with 10% DMSO/serum-free DMEM. After washing with serum-free media, 10 ml complete media was added. Conditioned media were sampled at 16 (hatched bars) and 24 h (black bars). Detection of mature pIL-1 β in the cell medium was accomplished by ELISA (R&D Systems). Media samples were diluted to achieve concentrations in the linear range of the ELISA assay (8–60 pg ml⁻¹). Background IL-1 β concentrations were



determined in media from cells transfected with the expression vector lacking ICE cDNA, and this value was subtracted from all other concentrations. The per cent activity values were calculated as the ratio of secreted IL-1 β from cells transfected with mutant ICE divided by IL-1 β secreted by cells transfected with wild-type ICE. Ratios are the means determined in at least two experiments.

present in cell lysates by western blotting using an anti-p20 rabbit antiserum, which also recognizes the intact p30 precursor. Similar amounts of p30 were observed for each mutant regardless of enzymatic activity, suggesting that even inactive mutants are efficiently transcribed, translated and folded (data not shown).

Inactive mutants such as Cys 285 to Ala and His 337 to Ala can be distinguished, however, from active mutants by the lack of a p20 species in the former, demonstrating that autoprocessing has been abrogated by these mutations.

Using the transient expression system, we determined that

FIG. 4 a, Electron density map of the active site. The $|2F_{\text{obs}} - F_{\text{calc}}|$ Fourier synthesis was calculated at 2.6 Å resolution with phases determined from the refined coordinates. The contour level is $+0.8\sigma$, where σ is the root-mean-square density throughout the unit cell. The refined coordinates of the final model are superimposed and identified by their three-letter codes and by sequence position. Carbon atoms are coloured white, nitrogen atoms blue, oxygen red and sulphur is yellow. Hydrogen bonds are depicted as dashed lines. **b**, Stereo stick drawing of the active site. Residues of ICE in contact with the tetrapeptide aldehyde inhibitor (coloured purple) are coloured by subunit, with p20 residues coloured green and p10 residues red. Residues of ICE that have at least one atom within 4 Å of any atom in the inhibitor are: Arg 179; Ser 236; His 237; Gly 238; Gln 283; Ala 284; Cys 285; Val 338; Ser 339; Trp 340; Arg 341; His 342; Pro 343; Met 345; Val 348; and Arg 383. The acetyl group of the inhibitor is adjacent to Pro 343. Definitions: P4, tyrosine; P3, valine; P2, alanine; P1, aspartic acid.

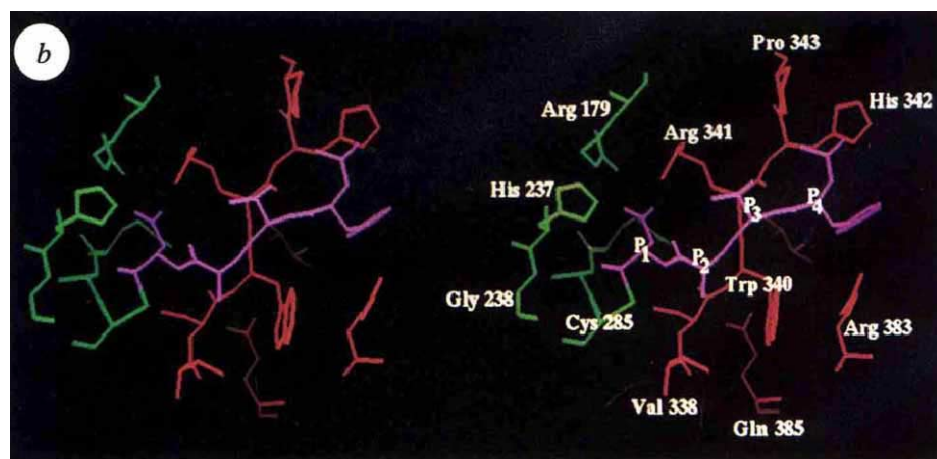
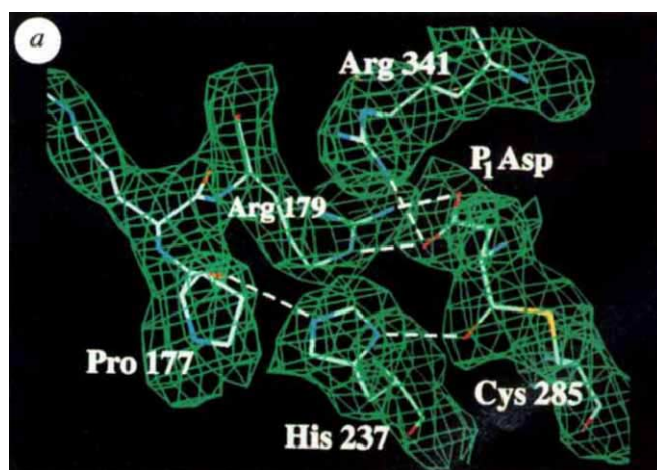


TABLE 1 Structure determination of ICE

Protein modification	Resolution (Å)	Completeness of data (%)	R_{merg}^* (%)	Unit cell dimensions (Å)		No. of sites	R_c^* (%)	Phasing power [*]
				a, b	c			
Tetrapeptide aldehyde [†]	20–2.6	90	8.3 (2)	64.4	163.3	—	—	—
Tetrapeptide aldehyde	20–2.8	78	8.3 (3)	64.7	162.9	—	—	—
Iodinated tetrapeptide aldehyde	20–3.5	86	9.4 (2)	64.4	162.8	2	0.88	1.09
Thimerosal	20–3.5	88	8.4 (1)	64.4	162.3	5	0.67	1.08
Gold thiomalate	20–3.5	74	9.5 (1)	64.7	162.7	3	0.72	1.22
Uranyl acetate	20–4.0	80	10.8 (2)	64.7	162.9	2	0.79	1.32
Lead chloride	20–3.5	64	8.9 (2)	64.7	162.8	2	0.76	1.38

Details of the expression, purification, autoproccessing, and sample preparation for crystallization will be published elsewhere (J.A.T. *et al.*, manuscript in preparation). Crystals of inhibited ICE were grown by vapour diffusion. Protein (20 mg ml⁻¹ in 50 mM citrate, 2.0 mM DTT, pH 6.5) was mixed with reservoir (15% PEG 4K, 400 mM LiSO₄, 200 mM sodium HEPES, 5 mM sodium cacodylate, 0.5% beta-octyl glucoside, pH 7.0) and allowed to stand over the reservoir solution at 4 °C. Crystals were equilibrated with 18% PEG 4K reservoir solution before data collection or derivatization. Refined heavy-atom parameters were used to compute multiple isomorphous replacement phases. The mean figure of merit was 0.65 to 3.5 Å resolution. Including the anomalous data for the Hg derivative showed the space group to be *P*₄₃₂₁ rather than its enantiomorph. The initial electron density map revealed the beta-sheet and five of the surrounding helices. Solvent flattening and phase extension³⁸ improved the map, and was followed by cycles of model building³⁹, positional refinement^{40,41}, and phase combination³⁸ until the switch to phases calculated from the model could be made. The current model is consistent with the derived amino-acid sequence of ICE^{7,8} and the chemical nature of the heavy-atom substitutions. The model has an *R*-factor of 19% against all observed data between 7 Å and 2.6 Å resolution, with root-mean-square deviation from ideal bond lengths and angles of 0.010 Å and 2.82°, respectively. We see no electron density for amino acids 120 to 130, and they are presumed disordered. We have confirmed their presence in the crystal by amino-acid sequence analysis. Atomic coordinates for ICE will be submitted to the Brookhaven Protein Data Bank.

* See ref. 42 for definitions.

† Data collected at -6 °C.

mutation of Cys 285 or His 237 eliminates pIL-1 β processing activity (Fig. 3) as well as autoproccessing, confirming their role in catalysis. Mutation of Arg 179, which contacts the P₁ Asp, to Glu also abolishes activity. Each of these active site residues is conserved in the human, murine and rat ICE complementary DNA sequences^{7,8,18,19}, as well as in three *ced-3* sequences⁹. Mutation of Cys 244 to Ala, which may contact P' side chains of substrates, reduces enzymatic activity significantly (Fig. 3). In contrast, mutation of other residues proximal to Cys 285 does not eliminate activity. Such residues include Ser 332, 333 or 339, and His 248, which are not conserved in *ced-3* sequences⁹.

The pH dependence of ICE enzymatic activity (k_{cat}/K_m) fits a sharp bell-shaped curve with a maximum at pH 7.5, with component pK_a s of 5.5 and 8.2. The basic limb of the pH dependence is consistent with protonation of a weak base such as His 237, which may function as a general acid-base during catalysis^{20,21} and/or increase binding affinity of substrate. The turnover number (k_{cat}) of ICE is low, about 1 s⁻¹ for the tetrapeptide *p*-nitroanilide substrate, compared to 42 s⁻¹ for papain for a dipeptidyl *p*-nitroanilide substrate²². ICE lacks a polarizing acidic residue in the active site, corresponding to papain Asp 158. This difference may help explain the low k_{cat} of ICE, which is comparable to those observed for Asp 158 mutants of papain²³.

ICE is a (p20)₂/(p10)₂ tetramer

ICE has been described as a heterodimer⁷, but the arrangement of p20 and p10 subunits in the crystal shows that ICE is a tetramer (Fig. 2). Two p20 subunits (green and blue) surround two adjacent p10 subunits (red and gold) that interact across the crystallographic 2-fold axis. The area of contact buried in this interface between p10 subunits is ~5,200 Å². In comparison, the surface area buried between subunits of proteins in the 24–40K range that form dimers is often ~1,800 Å² (ref. 24). This observation suggests that the crystallographic 2-fold axis corresponds to an oligomer interface solution, and that ICE is a (p20)₂/(p10)₂ tetramer.

Other evidence supports this hypothesis. Sequence comparison between human and murine ICE¹⁸ reveals that p10 is more highly conserved (81%) than p20 (59%). Much of the dimer-dimer interface consists of p10 residues 318–322 and 386–396. Thirteen of these sixteen residues are conserved, consistent with their location at the proposed tetramer interface. Four of the

six residues from the p20 subunits that make contacts across the 2-fold axis are also conserved. Mutation of His 322, a residue 35 Å away from the active site, to Ala, Lys or Gln eliminates ICE activity (Figs 1 and 3). Sensitivity to substitution at this residue is consistent with the higher association state we propose, where His 322 is completely buried in the interface between the p10 subunits. In solution, we have analysed the distribution of ICE (p20)_n/(p10)_n oligomers by size-exclusion chromatography and by quasi-elastic light scattering. A continuous distribution of ensembles up to very high molecular mass aggregates ($M_r > 10^6$), which may include a tetrameric species, is observed (data not shown).

The quaternary arrangement of p20/p10 subunits (Fig. 2) raises the question of which pair of subunits originates from a contiguous p45 precursor polypeptide. If we consider an ICE p20/p10 dimer (the green/red or the gold/blue pair in Fig. 2) to be derived from one p45 molecule, then the measured distance from the α -carbon of the p20 C-terminal residue (Asp 297) to the α -carbon of the N-terminal p10 residue (Ala 317) is 67 Å. This distance would have to be spanned by the 19 amino acids removed during processing of ICE to its active form. This possibility seems unlikely unless large conformational changes accompany autoproccessing. Alternatively, we propose that an ICE p20/p10 dimer is derived from two different p45 molecules, that is, the green/gold pair (or the blue/red pair) in Fig. 2 arises from the same p45 precursor. In this model, the distance between the α -carbons of Asp 297 and Ala 317 is 4.9 Å, and the residues preceding these form two anti-parallel β -strands that include six main-chain hydrogen bonds. A salt bridge between the p20 C terminus and the p10 N terminus is observed, and would be generated after p45 is processed to the active form.

Discussion

ICE defines a new class of cysteine proteases with respect to global fold, topology, and quaternary structure. The topology of ICE is different from cysteine proteases in the papain subfamily. The papain active site cysteine residue is located at the N terminus of an α -helix, whereas Cys 285 of ICE is at the C-terminal end of β -strand 4. ICE is more closely related to the recently discovered CED-3 protein⁹. Residues 246–320 of CED-3 share 43% sequence identity with residues 166–287 of human ICE⁹. This region of ICE includes all of the residues important for recognition of aspartic acid in P₁ (Arg 179, Gln 283) and for

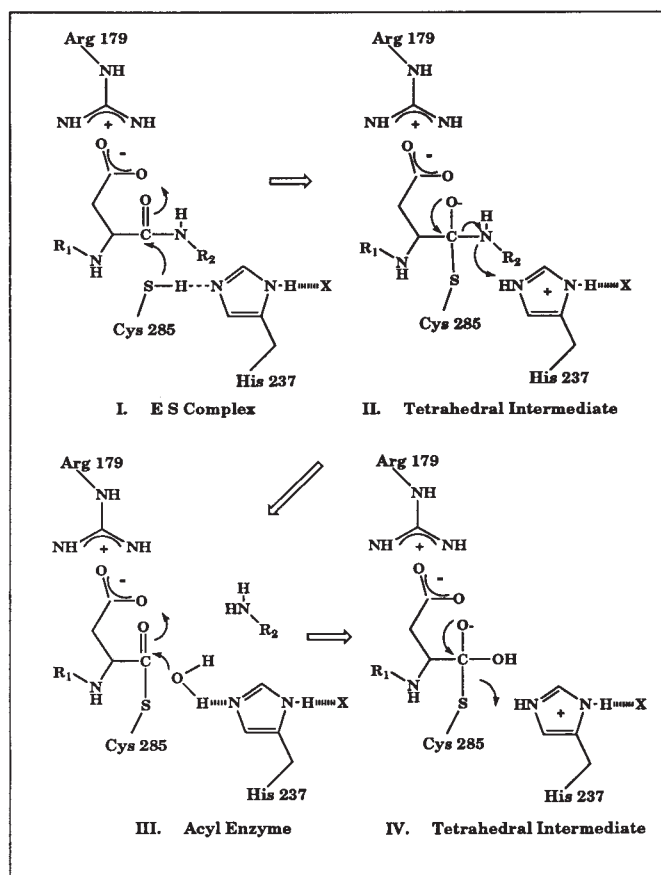


FIG. 5 A plausible mechanism for ICE catalysis of peptide hydrolysis using the Cys 285-His 237 catalytic diad. For the synthetic substrate used to quantify ICE activity, R_1 represents a Suc-Tyr-Val-Alanyl tripeptide and R_2 represents a *p*-nitrophenyl group. The structure of the peptide-aldehyde complex does not identify unambiguously a putative third residue (X) that would enhance the basicity of the His 237 imidazole ring by hydrogen bonding. The identity of this residue would vary depending on the torsion angles of the His 237 side chain in the enzyme-substrate complex. In the appropriate geometry, X might represent the carbonyl oxygen of Pro 177 or the P_1 carboxylate oxygen of substrate. The Arg 341 guanidino group may also be involved in electrostatic interactions with the substrate P_1 carboxylate.

catalysis (His 237 and Cys 285). These residues are all conserved in CED-3, suggesting that *ced-3* genes encode ICE-like cysteine proteases⁹.

The hydrolytic mechanism of ICE must proceed through formation of a substrate tetrahedral intermediate, probably requiring activation of the Cys 285 side chain by polarization and/or deprotonation by His 237 in a catalytic diad (Fig. 5). Two possible geometries for a substrate tetrahedral intermediate are suggested by the structure. In one, the oxyanion points away from the Cys 285 backbone amide and is in contact with the imidazole ring of His 237. This model is consistent with the structure reported here for the peptide aldehyde complex (Fig. 4a), but may be the opposite geometry of the enzyme-substrate tetrahedral intermediate. More likely, the oxyanion of the substrate tetrahedral intermediate interacts with the backbone amide protons of Cys 285 and Gly 238, which do not have hydrogen-bonding partners in the structure of the ICE-aldehyde complex. In this geometry, His 237 participates in general acid catalysis of acyl-enzyme formation by protonation of the tetrahedral intermediate aminol nitrogen atom (Fig. 5). Depending on the tor-

sion angles of the His 237 side chain, atoms such as the carbonyl oxygen of Pro 177, or perhaps the substrate P_1 carboxylate oxygen, may orient His 237 during catalysis.

We propose that the active form of ICE in solution is a tetramer of interdigitating subunits, a unique quaternary structure among known proteases. A tetrameric association state has implications for ICE autoproteolysis and regulation. In the tetramer, the $p10$ subunit from one ICE molecule complexes with the $p20$ subunit from a different molecule to create an active site. Two of the cleavage sites required to generate active ICE are located adjacent in space in this arrangement, and share a surface-exposed location that would aid processing.

ICE activity may be regulated allosterically because enzymatic activity is sensitive to mutations at the dimer-dimer interface, far removed from the active site (Figs 1 and 3). We have identified two non-competitive inhibitors of ICE, gold-thiomalate and auranofin, which are anti-inflammatory agents that are known to decrease circulating IL-1 β levels²⁵. Gold-thiomalate binds to Cys 364 and Cys 397, residues distant from the active site (~ 30 Å), and located adjacent to the 2-fold axis in the crystal (Table 1). Auranofin binds equally well to Cys 364 (data not shown). It is possible that the interface between ICE $p10$ subunits is a physiological target for these compounds.

The high-resolution structure of ICE we describe will accelerate efforts to design specific inhibitors of ICE. Elevated levels of IL-1 β have been detected in septic shock, rheumatoid arthritis, diabetes and other autoimmune diseases²⁶⁻³⁰. Inhibition of IL-1 β activity has been demonstrated to ameliorate symptoms in several animal models of human inflammatory disease³¹⁻³⁵. IL-1 β is also overexpressed in the brain of patients with Alzheimer's disease³⁶, and a recently reported clinical trial with Alzheimer's patients suggests that anti-inflammatory drugs may be preventative³⁷. ICE inhibitors may constitute a new type of anti-inflammatory agent, and ultimately find use in therapy of inflammation and neurodegenerative diseases. □

Received 4 March; accepted 23 June 1994.

- Black, R. A., Kronheim, S. R. & Sleath, P. R. *FEBS Lett.* **247**, 386-390 (1989).
- Kostura, M. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5227-5231 (1989).
- Sleath, P. R., Hendrickson, R. C., Kronheim, S. R., March, C. J. & Black, R. A. *J. biol. Chem.* **265**, 14526-14528 (1990).
- Howard, A. D. et al. *J. Immun.* **147**, 2964-2969 (1991).
- March, C. J. *Nature* **315**, 641-647 (1985).
- Singer, I. I. et al. *J. exp. Med.* **167**, 389-407 (1988).
- Thornberry, N. A. et al. *Nature* **356**, 768-774 (1992).
- Cerretti, D. P. et al. *Science* **256**, 97-100 (1992).
- Yuan, J., Shaham, S., Ledoux, S., Ellis, H. M. & Horvitz, H. R. *Cell* **75**, 641-652 (1993).
- Miura, M., Zhu, H., Rotello, R., Hartwig, E. A. & Yuan, J. *Cell* **75**, 653-660 (1993).
- Nett-Fioridali, M. A., Berson, D. R. & Chaplin, D. D. *J. Cell Biochem.* **17B**, 117 (1993).
- Ray, C. A. et al. *Cell* **69**, 597-604 (1992).
- Gagliardini, V. et al. *Science* **263**, 826-828 (1994).
- Marx, J. & Barinaga, M. *Science* **259**, 760-762 (1993).
- Chapman, K. T. *Biorg. med. Chem. Lett.* **2**, 613-618 (1992).
- Miller, D. K. et al. *Ann. NY Acad. Sci.* **696**, 133-148 (1993).
- Kamphuis, I. G., Drenth, J. & Baker, E. N. *J. molec. Biol.* **182**, 317-329 (1985).
- Nett, M. A. et al. *J. Immun.* **149**, 3254-3259 (1992).
- Shivers, B. D., Giegel, D. A. & Keane, K. M. *J. Cell Biochem.* **17B**, 119 (1993).
- Polgar, L. *Eur. J. Biochem.* **33**, 104-109 (1973).
- Baker, E. N. & Drenth, J. in *Biological Macromolecules and Assemblies* Vol. 3 (eds Jurnak, F. & McPherson, A.) 312-368 (Wiley, New York, 1987).
- Menard, R. et al. *Biochemistry* **30**, 8924-8928 (1991).
- Menard, R. et al. *Biochemistry* **30**, 5531-5538 (1991).
- Miller, S., Lesk, A. M., Janin, J. & Chothia, C. *Nature* **328**, 834-836 (1987).
- Bender, P. E. & Lee, J. C. A. *Rep. med. Chem.* **25**, 185-193 (1989).
- Dinarello, C. A. *Immun. Rev.* **127**, 119-146 (1992).
- Harris, E. D. *Jr New Engl. J. Med.* **322**, 1277-1289 (1990).
- Mandrup-Poulsen, T. et al. *Cytokine* **5**, 185-191 (1993).
- Hammerberg, C. et al. *J. clin. Invest.* **90**, 571-583 (1992).
- Noronha, I. L., Kruger, C., Andrassy, K., Ritz, E. & Waldherr, R. *Kidney Int.* **43**, 682-692 (1993).
- Dinarello, C. A., Gelfand, J. A. & Wolff, S. M. *J. Am. med. Ass.* **269**, 1829-1835 (1993).
- Geiger, T. et al. *Clin. exp. Rheumatol.* **11**, 515-522 (1993).
- Ohlsson, K., Björk, P., Bergenfeldt, M., Hageman, R. & Thompson, R. C. *Nature* **348**, 550-552 (1990).
- Woolley, P. H. et al. *Arthritis Rheumatol.* **36**, 1305-1314 (1993).
- Sandberg, J.-O., Eizirik, D. L., Sandler, S., Tracey, D. E. & Andersson, A. *Diabetes* **42**, 1845-1851 (1993).
- Griffin, W. S. T. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7611-7615 (1989).
- Breiter, J. C. S. et al. *Neurology* **44**, 227-232 (1994).
- CCP4-Collaborative Computing Project No. 4 A Suite of Programs for Protein Crystallography (Daresbury Laboratory, Warrington, 1979).
- Quanta version 4.0b (Molecular Simulations Inc., Burlington MA, 1992).

40. Brünger, A. T. *Acta crystallogr.* **A46**, 46–57 (1990).
 41. Brünger, A. T., Krukowski, A. & Erickson, J. *Acta crystallogr.* **A46**, 585–593 (1990).
 42. Blundell, T. L. & Johnson, L. N. *Protein Crystallography* (Academic, New York, 1976).
 43. Carson, M. J. *appl. Crystallogr.* **24**, 958–961 (1991).
 44. Eklund, H. *J. molec. Biol.* **102**, 27 (1976).
 45. Holmgren, A., Soderberg, B.-O., Eklund, H. & Branden, C.-I. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2305–2309 (1975).

ACKNOWLEDGEMENTS. We thank T. Fox and J. Coll for assistance with protein purification and chemical modification studies, M. Mullican, D. Lauffer and T. Tung for synthesis of the inhibitors used in the structure determination, O. Futer for assistance with mutagenesis, J. Fulghum for assistance with fermentation, M. Fleming for protein sequence analysis and oligonucleotide synthesis, G. Bemis and D. A. Pearlman for molecular dynamics analysis and discussions on mechanism, M. D. Sintchak for help with figures and J. Boger and V. Sato for reading of the manuscript.

LETTERS TO NATURE

Stars within the Large Magellanic Cloud as potential lenses for observed microlensing events

Kallash C. Sahu

Instituto de Astrofísica de Canarias, 38200 La Laguna, Tenerife, Spain

MASSIVE compact objects in the Galactic halo, known as MACHOs, have been postulated as the origin of a substantial fraction of the 'dark matter' known to exist in the haloes of galaxies^{1,2}. Paczyński³ has suggested that it might be possible to detect these low-luminosity objects by their potential to act as gravitational lenses, causing a characteristic brightening when they cross the path of light from a star in a nearby galaxy. Very recently, two groups reported possible detections of microlensing of stars in the Large Magellanic Cloud (LMC)^{4,5}, which was interpreted as a possible fingerprint of MACHOs. Here I show that microlensing by stars within the LMC itself can account for the observed events. In the future it should be possible to distinguish between the two possible sources of microlensing events, however, because events caused by stars in the LMC should be clustered toward the central region of that galaxy whereas those caused by MACHOs should be uniformly distributed over the whole LMC.

Based on a suggestion by Paczyński³, two groups (involved in the monitoring of a few million stars in the LMC for more than a year) have reported the detection of three microlensing events as evidence for the presence of dark objects in the Galactic halo^{4,5}. But this has also been the cause for some concern: if these events are caused by such objects in the halo, it would lead to some problems in terms of stellar evolution and the theory of galaxy formation⁶. Recent work on deuterium abundance derived from the observations of a quasar suggests that the dark haloes of galaxies may well be non-baryonic⁷. Furthermore, the rate of microlensing events as observed seems to be lower than expected for the dark halo, and it is important to estimate the rate expected from the stars that are known to exist in our Galaxy and in the LMC. The analysis for the Galactic stars has been done by Gould *et al.*⁸, but the importance of the LMC stars has been overlooked so far although Gould⁹ has investigated the possibility that MACHOs in the halo of the LMC might be the lenses. Here I present an estimate of the rate of microlensing events to be expected from the LMC stars acting as gravitational lenses.

To calculate the probability of the microlensing being caused by a star in the LMC itself, we need to know the stellar mass density there. First, I consider only the bar of the LMC. From the surface luminosity, it is estimated that the observed luminosity of the bar is ~10–12% of the total observed luminosity of the LMC in the optical wavelengths^{10,11}. To calculate the extinction, I have used the IRAS 100- μ m maps^{12,13}, the background galaxy counts¹⁴, and the observations of the most reddened stars in the region¹⁵, which together give a consistent value of 1.5 mag

in the V band. To see how the extinction affects the mass, I assume that the extinction is uniform in depth. Thus for any line of sight, if A_v is the total extinction in the line of sight, then the ratio of the observed to true luminosity at any point

$$\frac{L_{\text{obs}}}{L_{\text{true}}} = \frac{1}{d} \int_0^d 2.5^{-A_v/d} dl = \frac{1 - e^{-0.916A_v}}{0.916A_v} \quad (1)$$

where d is the total depth and l corresponds to the direction of the line of sight. Using $A_v = 1.5$ mag for the bar, and $A_v = 0.3$ – 0.4 mag for the region outside, it is easy to show that the true luminosity of the bar is ~14–18% of the total luminosity of the LMC. The total mass of the LMC, as determined from various methods, is in the range $(6\text{--}15) \times 10^9 M_\odot$ (refs 10, 16) where $M_\odot$ is the solar mass. Assuming that the mass-to-light ratio in the bar and the outer parts are the same, the mass of the bar can be taken as $2 \times 10^9 M_\odot$. Considering the fact that the LMC is gas-poor and only ~5% of the LMC mass is neutral hydrogen^{16–18}, I will neglect the contribution of gas to the mass and assume that the entire mass is made up of stars.

To calculate the probability of microlensing, I assume that there are N_{tot} stars being monitored. To determine the effect of extinction on the number of monitored stars at various depths, I note that the limiting magnitude of the current surveys are ~20–21, which at the distance of the LMC, corresponds to an absolute magnitude of 1.5–2.5. So the magnitude of the stars that can be observed at the near side is 1.5 to 2.5, whereas the limiting magnitude at the far end is 1.5 mag brighter. If the distribution of stars among different spectral types is assumed to be similar to that observed in the solar neighbourhood, then the difference in the monitored number of stars per unit depth from front end to the back is ~3 (ref. 19). (As it turns out, the effect of extinction is small. If the extinction is 3 instead of 1.5 mag, the net probability decreases only by <50%.)

Hence assuming the extinction to be uniform in depth, the observed number of stars at any layer dl , at a depth of l , can be expressed as

$$N_{\text{obs}}(l) = N_1 3^{-l/d} dl \quad (2)$$

where N_1 is the observed number of stars per unit depth in absence of extinction. If N_{tot} is the total number of stars being monitored, then

$$N_{\text{tot}} = \int_0^d N_1 3^{-l/d} dl = 0.6 N_1 d \quad (3)$$

Substituting equation (3) in equation (2)

$$N_{\text{obs}}(l) \approx \frac{N_{\text{tot}}}{0.6d} 3^{-l/d} dl \quad (4)$$

The fraction of area covered by the Einstein rings of all the individuals stars lying in the front of this layer can be expressed as

$$A_f(l) = \int_0^l \pi R_E^2(l) n(l) dl \quad (5)$$

where $n(l)$ is the stellar number density at depth l and R_E is the Einstein radius. Note that $n = \rho/m$ and $R_E^2 \propto m$ where ρ is the

stellar mass density and m is the mass of the star. Assuming the stellar density to be uniform with depth

$$A_f(l) = \frac{2\pi G \rho l^2}{c^2} \quad (6)$$

which is thus independent of the mass distribution. From equations (4)–(6), the optical depth to microlensing by stars within the bar can be expressed as

$$\tau_{\text{bar}} = \frac{1}{N_{\text{tot}}} \int_0^d N_{\text{obs}}(l) A_f(l) dl \approx \frac{0.5\pi G \rho_{\text{bar}} d^2}{c^2} \quad (7)$$

The value of ρ can be substituted by

$$\rho_{\text{bar}} = \frac{M}{LWd} \quad (8)$$

where L , W , d and M are the length, width, depth and mass of the bar respectively. Assuming $L = 3,000$ pc and $W = d = 600$ pc (refs 10, 20, 21), and substituting equation (8) in equation (7), we get

$$\tau_{\text{bar}} \approx 5 \times 10^{-8} \quad (9)$$

with some uncertainties due to the uncertainties in the size and mass of the bar. Carrying out an identical analysis for the region outside the bar it is easy to see that, if the depth of the LMC disk is 100–300 pc (ref. 22), the optical depth in the region outside the bar is ~ 4 –12 times smaller.

The MACHO⁴ event lies in the central region of the bar. Keeping in mind the uncertainties involved in the estimation of optical depth on the basis of a single event^{8,23}, we can only say that it seems to be well below the optical depth of 5×10^{-7} expected from a dark halo made up entirely of MACHOs, and is consistent with the optical depth calculated above. In the case of EROS⁵ events, one lies in the outer region of the bar, the other is far from the bar, and the optical depth has been estimated to be higher ($\sim 2 \times 10^{-7}$; ref. 8). But Gould *et al.*⁸ also suggest that there may be some systematic effects and the detection efficiencies may be uncertain. The situation will be clearer as more events are observed, particularly those with higher magnifications (see below), and we must wait for more events to be observed before making any direct comparison with the calculated optical depth.

If the microlensing is indeed caused by the LMC lenses, then one important consequence is that the sources can be resolved at much smaller magnifications, as shown below.

The amplification rises sharply when the lensing object comes close to being perfectly aligned with the source. The physical reason for this is that when the lensing object and the source are perfectly aligned, the image becomes a ring (also called the Einstein ring) instead of two separate images. (For details, see ref. 3.) In the case of lensing by MACHOs, R_E is of the order of $(1.2 \times 10^{14})(M/M_\odot)^{1/2}$ cm, which, at the source plane, is $\sim 10^{15}(M/M_\odot)^{1/2}$ cm. If the lensing is caused by the objects in the disk of our galaxy, R_E projected onto the source plane is $> 10^{15}(M/M_\odot)^{1/2}$. But if the lens is in the LMC, for a typical distance D of ~ 100 pc between the source and the lens, $R_E \approx 10^{13}(M/M_\odot)^{1/2}$ cm, which is at least two orders of magnitude smaller. As an example, consider a red giant source, and a $0.5M_\odot$ lens. In order that at least a part of the source is perfectly aligned with the lensing object in this case, the impact parameter u has to be $\lesssim 0.001$ for a halo or disk lens whereas it has to be $\lesssim 0.1$ for an LMC lens, which will make the lightcurve different from that of a point source²⁴. Considering the fact that the events are expected to be uniformly distributed in u , this provides a possible way to distinguish between the halo lenses and the LMC lenses, as more events are observed.

As is clear from the optical-depth estimates, the LMC-induced events should be strongly concentrated towards the central region of the LMC, whereas the frequency of galactic events

(whether by disk or halo lenses) should be proportional to the density of stars observed. $\square$

Received 8 March; accepted 4 July 1994.

1. Begeman, K. thesis, Groningen Univ. (1987).
2. Fich, M., Blitz, L. & Stark, A. *Astrophys. J.* **342**, 272–284 (1989).
3. Paczyński, B. *Astrophys. J.* **304**, 1–5 (1986).
4. Alcock, C. *et al. Nature* **365**, 621–623 (1993).
5. Aubourg, E. *et al. Nature* **365**, 623–625 (1993).
6. Hogan, C. J. *Nature* **365**, 602–603 (1993).
7. Songaila, L. L., Cowie, C. J., Hogan, C. J. & Rugers, M. *Nature* **368**, 599–604 (1994).
8. Gould, A., Miralda-Escudé, J. & Bahcall, J. N. *Astrophys. J.* **423**, L105–L108 (1994).
9. Gould, A. *Astrophys. J.* **404**, 451–454 (1994).
10. De Vaucouleurs, G. & Freeman, K. C. *Vistas Astr.* **14**, 163–294 (1973).
11. Bothum, G. D. & Thompson, I. B. *Astr. J.* **96**, 877–883 (1988).
12. Schwing, P. B. W. & Israel, F. P. *Atlas and Catalog of Infrared Sources in Magellanic Clouds* (Kluwer Academic, Dordrecht, 1990).
13. Laureijs, R. thesis, Groningen Univ. (1989).
14. Hodge, P. in *The Magellanic Clouds* (eds Haynes, R. & Milne, D.) 57–62 (Int. Astr. Un. Symp. No. 148, Kluwer Academic, Dordrecht, 1991).
15. Isserstedt, J. *Astr. Astrophys.* **41**, 175–182 (1975).
16. Westerlund, B. *Astr. Astrophys. Rev.* **2**, 29–78 (1990).
17. Israel, F. P. & de Graauw, Th. in *The Magellanic Clouds* (eds Haynes, R. & Milne, D.) 45–52 (Int. Astr. Un. Symp. No. 148, Kluwer Academic, Dordrecht, 1991).
18. Rohlf, K., Kreitschmann, J., Siegman, B. C. & Feitzinger, J. V. *Astr. Astrophys.* **137**, 343–357 (1984).
19. Allen, C. W. *Astrophysical Quantities* 3rd edn (Athlone, London, 1973).
20. Binney, J. & Tremaine, S. *Galactic Dynamics* (Princeton Univ. Press, 1987).
21. Westerlund, B. in *The Magellanic Clouds* (eds Haynes, R. & Milne, D.) 15–23 (Int. Astr. Un. Symp. No. 148, Kluwer Academic, Dordrecht, 1991).
22. Feast, M. in *Recent Developments in Magellanic Cloud Research* (eds de Boer, K. S., Spite, F. & Stasinska, G.) 75–86 (Observatoire de Paris, 1989).
23. Gould, A. *Astrophys. J.* **421**, L71–L74 (1994).
24. Sahu, K. C. *Publ. Astro. Soc. Pacific* (in the press).

ACKNOWLEDGEMENTS. H. Butcher, P. Hodge and B. Westerlund helped to clarify some doubts about the structure of the LMC. Some details of the observational program provided by K. Freeman are gratefully acknowledged. I thank the referees B. Paczyński and Alain Milstajn for their comments, which helped to improve the manuscript.

Coupling between adjacent crystal planes in heterogeneous catalysis by propagating reaction–diffusion waves

V. Gorodetskii, J. Lauterbach, H.-H. Rotermund, J. H. Block & G. Ertl

Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4–6, D-14195 Berlin, Germany

UNDERSTANDING of the mechanisms and kinetics of heterogeneous catalytic reactions has come largely from the study of gas–solid interactions on well defined single-crystal surfaces^{1,2}. But real catalysts usually consist of nanometre-sized particles on which several different crystal planes are exposed. In general, it has been assumed that their properties can be regarded as a superposition of the contributions from each individual structural element. Here we show that this assumption may be invalid, even qualitatively, in certain cases. We have studied the oxidation of hydrogen on platinum surfaces at low pressure and room temperature. On a macroscopic Pt(100) single crystal the reaction reaches a steady state with a uniform distribution of adsorbates. But on the platinum tip of a field ion microscope, on which several different crystal planes are exposed, the reaction has a very different character. The tip contains a region of the (100) plane just 40 nm in diameter, on which the reaction rate displays sustained temporal oscillations. This effect is associated with continuously changing distributions of the adsorbed species in the form of propagating waves, which are generated by coupling of reactions occurring on adjacent crystal planes. This kind of interaction between different crystal planes may exert a profound influence on the kinetics of heterogeneous catalysis.

The rate of a reaction catalysed by a uniform (that is, single-crystal) surface is usually formulated in terms of the surface concentrations (the partial coverages) of the reacting species and the associated rate constants. The associated set of coupled nonlinear differential equations generally has stationary solutions, but often also produces spatiotemporal patterns. Such behaviour has been verified experimentally with various systems^{3,4}. The kinetics may become oscillatory or even chaotic, and the lateral distributions of the adsorbed species are no longer uniform, but may exhibit various types of pattern formation on mesoscopic length scales, typically of the order of 1 μm , as determined by coupling of surface diffusion and reaction.

Our studies with the macroscopic (~ 10 mm diameter) Pt(100) single-crystal surface were performed with an ultra high vacuum apparatus which was equipped with a photoemission electron microscope (PEEM)⁵. The PEEM technique enables *in situ* imaging of adsorbate concentration distributions with a lateral resolution of ~ 0.2 μm and a temporal resolution of 20 ms. Its contrast is based on the variation of the (local) work function associated with the formation of chemisorption complexes. With the present system, adsorbed oxygen atoms increase the work function much more than do adsorbed hydrogen atoms^{6,7}. Therefore predominantly oxygen-covered areas appear darker in the PEEM images than those parts of the surface on which chemisorption of hydrogen prevails. Chemisorbed hydrogen and oxygen atoms react with each other far below room temperature, whereas water desorbs at ~ 180 K (refs 8, 9). Thus at the temperatures applied here (≥ 300 K), when a water molecule forms, it will be present on the surface only as a short-lived transient species.

Experiments with the Pt(100) sample were performed at temperatures between 300 and 400 K, and with varying H_2 and O_2 partial pressures up to 10^{-3} mbar. No sustained concentration patterns were observed: rather, the surface (and hence also the reaction rate) always reached a steady state with uniform adsorbate distribution as monitored by PEEM. If the temperature and one of the partial pressures were kept constant while the other one was varied, a gradual change of the brightness of the PEEM image signalled the continuous transition from pure H coverage to practically pure O coverage or *vice versa*. Closer inspection, however, revealed that this transition is not smooth but exhibits discontinuous features known as kinetic phase transitions^{10,11}: changing one of the partial pressures in small steps causes, within a critical range, a discontinuous switching of the state of the surface. As an example, Fig. 1 shows a series of PEEM images recorded at 300 K at 60-s intervals with a constant partial pressure of hydrogen $p_{\text{H}_2} = 5 \times 10^{-6}$ mbar. On lowering the O_2 partial pressure p_{O_2} to $\sim 6.4 \times 10^{-4}$ mbar, brighter H-rich patches nucleate and continuously grow in the darker O-rich surroundings, until the whole surface area is again uniform. It should be noted that p_{O_2} has to be about two orders of magnitude higher than p_{H_2} in order to achieve comparable coverages by adsorbed oxygen and adsorbed hydrogen, O_{ad} and H_{ad} , respectively. This is due to the fact that the sticking coefficient for dissociative oxygen adsorption is much lower than that for hydrogen¹² on a Pt(100) surface that has undergone reconstruction to a hexagonal structure. In the opposite direction (increasing p_{O_2}), the analogous transition occurs at somewhat higher $p_{\text{O}_2} \approx 8.0 \times 10^{-4}$ mbar, indicating the occurrence of hysteresis. For a given set of parameters, the concentration waves propagate on mesoscopic length scales at constant speed, ranging from ~ 5 $\mu\text{m s}^{-1}$ to 100 $\mu\text{m s}^{-1}$ in the parameter range indicated. Interestingly, the transition $\text{H} \rightarrow \text{O}$ was always found to be considerably faster than that in the reverse direction. These wave propagation speeds were the higher the more the adjusted partial pressure exceeded its critical value for the kinetic phase transition. It has, however, to be emphasized that these effects are only of a transient nature, and lead to steady-state behaviour with uniform concentration distributions.

These findings can be rationalized in terms of a model

developed for the kinetics of this system under steady-state flow conditions by Zhdanov and Kasemo¹³. This model is based on a mechanism which involves dissociative chemisorption of oxygen and hydrogen followed by the recombination steps $\text{H}_{\text{ad}} + \text{O}_{\text{ad}} \rightleftharpoons \text{OH}_{\text{ad}}$, $\text{H}_{\text{ad}} + \text{OH}_{\text{ad}} \rightarrow \text{H}_2\text{O}$. By using reasonable kinetic parameters for the individual reaction steps as supported by experimental evidence¹⁴, solution of the corresponding differential equations shows that at fixed temperature T , for small $p_{\text{H}_2}/p_{\text{O}_2}$ the surface is predominantly covered by oxygen, whereas for large $p_{\text{H}_2}/p_{\text{O}_2}$ it is mainly covered by hydrogen. In an intermediate region the system becomes bistable, with a hysteresis extending over a certain range of $p_{\text{H}_2}/p_{\text{O}_2}$ which becomes narrower for lower total pressures. Transition from one state to the other in such a homogeneous bistable reaction-diffusion system is characterized by nucleation and growth of concentration waves (just as observed here) known as 'chemical waves'¹⁵. For a perfectly homogeneous medium, a critical size of the nucleus would be needed for its further growth. This critical size requirement may, however, be circumvented by the presence of defects on the surface with modified kinetic parameters which may act as 'triggers'.

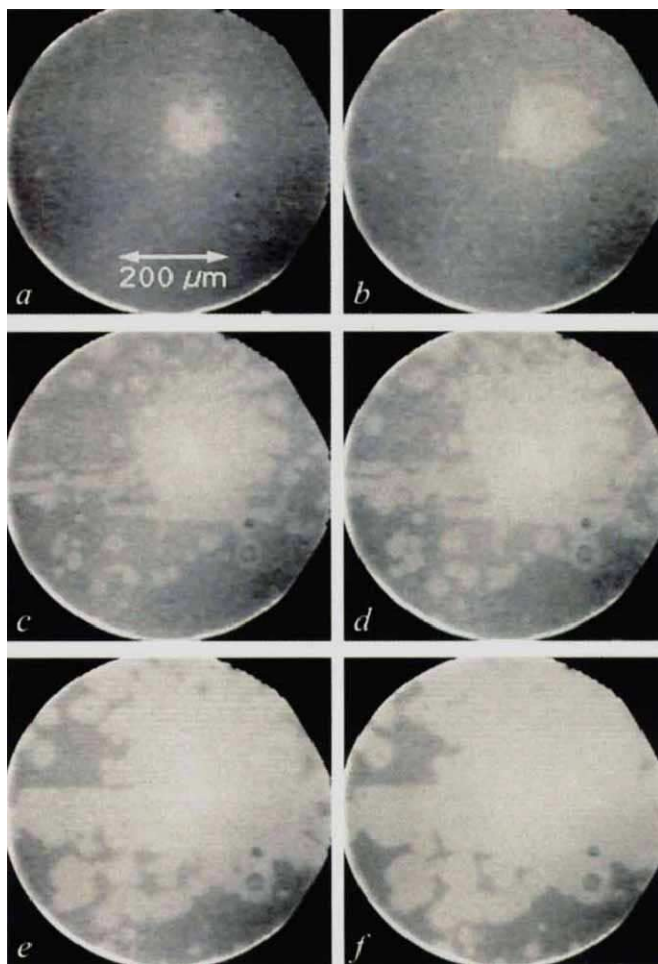


FIG. 1. a–f, Sequence of photoemission electron microscopy (PEEM) images from a macroscopic Pt(100) single-crystal surface under conditions of catalytic oxidation of hydrogen, showing transient nucleation and propagation of concentration waves on a mesoscopic scale after adjustment of the partial pressure of H_2 to a constant value ($p_{\text{H}_2} = 10^{-6}$ mbar), while the two other control parameters are kept fixed ($p_{\text{O}_2} = 6.4 \times 10^{-4}$ mbar, $T = 300$ K). The bright areas are mainly covered by chemisorbed H atoms, whereas the dark areas are essentially O-covered. (Time interval between subsequent images, 60 s; diameter of images, 0.6 mm).

Adjacent parts of the surface with different crystallographic orientation may act as such nucleation centres. We investigated this possibility with a field emitter tip. This tip exposes adjacent crystal planes with typical dimensions of the order of tens of nanometres, similarly to the situation found with the small metal particles of 'real' catalysts. Corresponding experiments were performed in an ultra high vacuum apparatus which was equipped with a sensitive channelplate for intensification of field ion microscopy (FIM) images. The surface of the Pt tip was cleaned by field evaporation. It was found in a previous study¹⁶ that instead of the usual noble-gas imaging technique, O_2 may be used for this purpose, producing images whose intensities were found to be higher for O_{ad} -covered than for H_{ad} -covered surfaces. In the present work it was found that even brighter images are produced by H_2O^+ ions being formed from the desorbing reaction product, so that in a unique fashion the momentarily catalytically active sites could be identified. It should be emphasized that the applied field strength of $1.5 \times 10^8 \text{ V cm}^{-1}$ was without any noticeable influence on the chemical processes on the surface, even if it was sufficient to ionise molecules in front of the surface. Pulsing, or even reversing, the polarity of the electric field did not affect the propagation velocities of concentration waves. In addition, recent quantitative studies revealed that even fields of up to $3 \times 10^8 \text{ V cm}^{-1}$ are without marked influence on the binding energy of hydrogen¹⁷.

Figure 2 shows a sequence of FIM images recorded at 300 K, $p_{O_2} = 5 \times 10^{-4}$ mbar and $p_{H_2} = 6 \times 10^{-4}$ mbar. As there is some uncertainty about the determination of effective partial pressures in FIM experiments, one should not compare the absolute values in both types of measurements, but rather notice that in the present case the ratio $p_{O_2} : p_{H_2}$ is much smaller than in the experiment shown in Fig. 1. The extended Pt(100) single-crystal surface would under these conditions be in a steady state, and predominantly uniformly covered by H_{ad} due to the much lower sticking coefficient for O_2 than for H_2 (apart from the fact that the impingement rate for H_2 is higher than that for O_2 by a factor of 4 at equal partial pressures)¹⁸. By contrast, the (100) plane with $\sim 40 \text{ nm}$ diameter in the central region of the FIM tip may be 'fed' by chemical waves from adjacent crystal planes. Under proper conditions (such as applied here) this may even give rise to a periodic switching between states with high hydrogen and oxygen coverage, respectively, and so lead to temporal oscillations of the reaction rate.

This is demonstrated by the series of FIM images shown in Fig. 2. Their different degrees of brightness characterize various states of the surface. Medium brightness is due to largely O_{ad} -covered areas (imaged by O_2^+ ions). Very bright zones indicate a high yield of H_2O ; these molecules are ionized after desorption. The sequence of images in Fig. 2 is interpreted as follows: the oscillation cycle starts with a largely O-covered surface, on which the hexagonal restructuring¹⁹ has been eliminated. In the region of the (100) plane, that is, near the centre of the image, a zone of high reactivity (which appears brighter due to imaging by H_2O^+ ions) nucleates and rapidly spreads across the whole surface (Fig. 2b–e). Due to the high reactivity of the Pt(100)- 1×1 surface at this stage, the coverages of both reacting species become rather low. As a consequence, the 1×1 -surface is no longer stable but transforms into the hexagonal phase (Fig. 2f, g). This process manifests itself by the appearance of fluctuating clusters of Pt atoms which can be discerned in the original video recording. Because the sticking coefficient for O_2 is much smaller on the Pt(100)-hex surface, parallel to this structural transformation the build-up of a H_{ad} -layer takes place whereby the hexagonal reconstruction is only partly eliminated²⁰. This would be the steady-state situation for an extended Pt(100)-surface under these conditions. In the present case, however, the adjacent (331)-type planes still exhibit a high oxygen sticking coefficient, and provide sources for waves propagating (and reacting) towards the central H_{ad} -region whereby simultaneously the transformation of the structure of the (100)-zone back to 1×1 is completed (Fig. 2i–l). The H_{ad} -regions appear somewhat darker in the FIM images, and hence their shrinking diameter becomes evident in Fig. 2k, l, until the surface is completely covered by O_{ad} . This represents the situation of Fig. 2a; the oscillating cycle is completed.

In this way the formation of H_2O molecules indeed shows sustained temporal oscillations, which means that (in contrast to the behaviour of macroscopic single-crystal surfaces), the kinetics are no longer stationary. The periodic variation of the state of the surface, accompanied by temporal oscillations of the integral reaction rate, is restricted to a relatively narrow range of external control parameters. Of decisive importance is the coupling of the (100) plane to adjacent different crystal planes, leading to propagating concentration waves and to modification of the reactivity. Occurrence of these effects is restricted to small samples such as the FIM tip, and will not happen on macro-

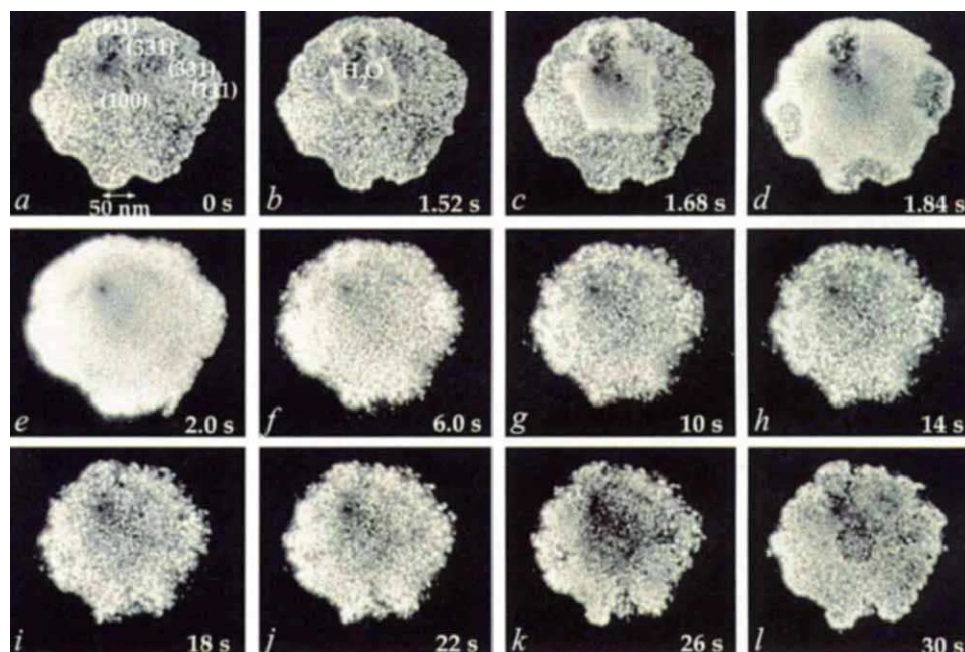


FIG. 2. a–l, Sequence of field ion microscopy (FIM) images from a Pt tip with 180 nm radius under constant conditions of catalytic oxidation of hydrogen; $p_{H_2} = 6 \times 10^{-4}$ mbar, $p_{O_2} = 5 \times 10^{-4}$ mbar and $T = 300 \text{ K}$. Field strength is $1.5 \times 10^8 \text{ V cm}^{-1}$; imaging gases are O_2 and H_2O formed by the reaction. The central region shows (100) orientation. (The time after the initial image (a) is shown at the lower right of each image.)

scopic single-crystal planes where waves triggered from the boundaries will die out within short ($\sim 1 \mu\text{m}$) distances.

It has to be emphasized that the phenomena reported here are fundamentally different from the concepts of 'spillover' or 'remote control' as frequently discussed in heterogeneous catalysis, where reacting species are supplied to the active centres by diffusion from other (more inactive) parts of the surface. With the present system, chemisorbed oxygen atoms would, at 300 K, not be mobile enough to account for the observed changes of the surface concentrations. By contrast it is the local coupling

between diffusion and reaction that gives rise to rapid propagation of reaction fronts. The propagation speed of such 'chemical waves' is, to a first approximation, determined by the product of a diffusion coefficient with an effective rate constant^{15,21,22}.

The present findings are of essential relevance to understanding heterogeneous catalysis. They indicate that the catalytic properties of a small particle exhibiting various crystal planes of restricted dimensions (typically $< 10 \text{ nm}$) cannot simply be regarded as a superposition of contributions from its individual parts. $\square$

Received 30 December 1993; accepted 14 June 1994.

- Somorjai, G. A. *Surf. Sci.* **229/300**, 849–866 (1994).
- Ertl, G. *Surf. Sci.* **229/300**, 742–754 (1994).
- Ertl, G. *Science* **254**, 1750–1755 (1991).
- Imbihl, R. *Progr. Surf. Sci.* **44**, 185–344 (1993).
- Engel, W., Kordes, M. E., Rotermund, H.-H., Kubala, S. & von Oertzen, A. *Ultramicroscopy* **36**, 148–153 (1991).
- Rotermund, H.-H., Jakubith, S., von Oertzen, A. & Ertl, G. *J. chem. Phys.* **91**, 4942–4948 (1989).
- Pennemann, B., Oster, K. & Wandelt, K. *Surf. Sci.* **249**, 35–43 (1991).
- Norton, P. R. in *The Chemical Physics of Solid Surfaces and Heterogeneous Catalysis Vol. 4* (eds King, D. A. & Woodruff, D. P.) 27–72 (Elsevier, Amsterdam, 1982).
- Kwasniewski, V. J. & Schmidt, L. D. *J. phys. Chem.* **96**, 5931–5938 (1992).
- Bykov, V. I., Elokhin, V. I., Gorban, A. N. & Yablonski, G. S. *Kinetic Models of Catalytic*

- Reactions*. (Elsevier, Amsterdam, 1991).
- Evans, J. W. *Langmuir* **7**, 2514–2519 (1991).
 - Guo, X. C., Bradley, J. M., Hopkinson, A. & King, D. A. *Surf. Sci.* **292**, L786–L790 (1993).
 - Zhdanov, V. P. & Kasemo, B. *Surf. Sci. Rep.* **20**, 111–189 (1994).
 - Hellsing, B., Kasemo, B., Zhdanov, V. P. *J. Catal.* **132**, 210–228 (1991).
 - Mikhailov, A. S. *Foundations of Synergetics Vol. 1* (Springer, Berlin, 1990).
 - Gorodetski, V., Drachsel, W. & Block, J. H. *Catal. Lett.* **19**, 223–231 (1993).
 - Ernst, N., Block, J. H., Kreuzer, H. J. & Ye, X. *Phys. Rev. Lett.* **71**, 891–894 (1993).
 - Ljungström, S., Kasemo, B., Rosén, A., Wahnström, T. & Fridell, E. *Surf. Sci.* **216**, 63–92 (1989).
 - Bartean, M. A., Ko, E. I. & Madix, R. J. *Surf. Sci.* **102**, 99–117 (1981).
 - Norton, P. R., Davies, J. A., Creber, D. K., Sitter, C. W. & Jackman, T. E. *Surf. Sci.* **108**, 205–224 (1981).
 - Luther, R. Z. *Elektrochem.* **12**, 596–598 (1906).
 - Arnold, R., Showalter, K. & Tyson, J. J. *chem. Educ.* **64**, 740–744 (1987).

Counting polymers moving through a single ion channel

Sergey M. Bezrukov*†, Igor Vodyanoy*‡ & V. Adrian Parsegian§

* Division of Intramural Research, NIDDK, § Laboratory of Structural Biology, DCRT, National Institutes of Health, Bethesda, Maryland 20892, USA

† St Petersburg Nuclear Physics Institute of the Russian Academy of Sciences, Russia 188350

‡ Office of Naval Research, Arlington, Virginia 22217-5000, USA

THE change in conductance of a small electrolyte-filled capillary owing to the passage of sub-micrometre-sized particles has long been used for particle counting and sizing. A commercial device for such measurements, the Coulter counter, is able to detect particles of sizes down to several tenths of a micrometre^{1–3}. Nucleopore technology (in which pores are etched particle tracks) has extended the lower limit of size detection to 60-nm particles by using a capillary of diameter $0.45 \mu\text{m}$ (ref. 4). Here we show that natural channel-forming peptides incorporated into a bilayer lipid membrane can be used to detect the passage of single molecules with gyration radii as small as $5\text{--}15 \text{ \AA}$. From our experiments with alamethicin pores we infer both the average number and the diffusion coefficients of poly(ethylene glycol) molecules in the pore. Our approach provides a means of observing the statistics and mechanics of flexible polymers moving within the confines of precisely defined single-molecule structures.

The underlying idea of our experiment is very similar to the resistive pulse principle that has been used in Coulter counters since 1953. If a nonconducting particle suspended in a conducting medium moves into a small capillary, it decreases the conductance of the capillary relative to that of the capillary filled with the conducting medium alone. The magnitude of this decrease in conductance is related to particle size (ref. 3).

For a molecular pore of $\sim 50 \text{ \AA}$ length and $\sim 10 \text{ \AA}$ radius (similar to alamethicin pore dimensions^{5,6}), the brownian motion of particles dominates over directional flow at all reasonable hydrostatic pressure differences across the membrane. Estimates for flow transient time and diffusion relaxation time (see below) show that for $5\text{-\AA}$ -radius particles in water these times become equal only at hydrostatic pressure differences of $\sim 10^8 \text{ Pa}$, equivalent to a 10-km-high water column.

To study diffusion-driven exchange of poly(ethylene glycol)s (PEGs) of different relative molecular masses (M_r) in a molecular pore, we analysed polymer-induced conductance fluctuations of a single alamethicin channel⁷ in the experiment schematically shown in Fig. 1. Addition of polymers to the membrane-bathing solution changes the conductance of the pore-containing membrane: the amplitudes of conductance states decrease and the current noise of the open channel increases. Figure 2 illustrates these changes for three states of the alamethicin channel, comparing spontaneous current "bursts" in polymer-free solution with those recorded in solutions of 15 wt% PEG with M_r 600 (PEG 600).

Figure 3 shows that the magnitude of polymer-generated excess noise is different for different conductance levels and

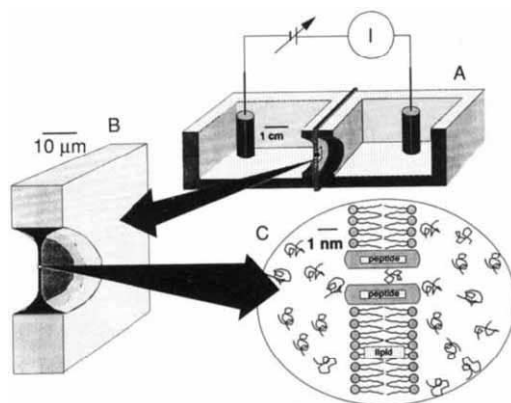


FIG. 1 Schematic representation of the cell (A), Teflon partition with a lipid bilayer (B), and a single ion pore exchanging polymers with bulk solution (C). Bilayer lipid membranes were prepared as described by Montal and Mueller¹⁸. The membrane-forming solution was α -diphytanoyl lecithin in n -pentane. Hexadecane in n -pentane (1:10) was used for aperture pre-treatment. Natural alamethicin, purified as described elsewhere¹⁹, was added as an ethanolic solution (after membrane formation) to the bathing solution from one side of the membrane only. Single-channel spontaneous insertions were detected as current 'bursts' shown in Fig. 2. PEGs of different M_r were added at 15% w/w concentration to aqueous 1 M NaCl solutions buffered at pH 6.2 by MES. All measurements were made at room temperature, at a transmembrane voltage of +150 mV on the side of peptide addition.

depends on polymer M_r . Noise is not monotonic in polymer size. Rather, there is a clear maximum for M_r 600–1,000. Small polymers (M_r 200–400) produce relatively less noise, presumably because their effect is an average over a large number of small events. Large polymers ($M_r > 1,000$) are geometrically restricted from even entering the channel.

The average number of polymer molecules in the channel, $\langle N(w) \rangle$, can be determined from the degree of channel conductance reduction on addition of polymer^{5,8}. If the entry of a single polymer of M_r w lowers single-channel conductance by an amount $h(w)$, and if one is working at a level of occupancy $\langle N(w) \rangle$ such that $h(w)\langle N(w) \rangle$ is much smaller than the mean conductance of polymer-free channel, one can write

$$h(w)\langle N(w) \rangle = h_{ch}(\infty) - h_{ch}(w)$$

Here $h_{ch}(w)$ is the mean conductance of the channel in the presence of polymers of M_r w , and $h_{ch}(\infty)$ its conductance in the presence of totally excluded, large polymers. These average conductances are measured directly (see Fig. 2).

We have verified previously⁵ that the relative reduction of alamethicin channel conductance in the presence of small, easily penetrating PEG 200 is close to the relative reduction in the conductivity of bulk electrolyte solutions on polymer addition. This was also shown for another mesoscopic channel formed by staphylococcal α -toxin⁹. Taking into account the large size of these channels, this means that in the case of small polymers their incremental (negative) contribution to channel conductance is close to the contribution that one would expect for the conductance of an electrolyte layer whose thickness is equal to the channel length. This equivalence permits a straightforward derivation of $h(w)$ from the polymer-induced reduction in solution conductivity. As polymer size is increased, the channel starts to exclude polymers; the proportional effect of polymer addition on channel conductance decreases⁵. As long as the size of a polymer coil does not significantly exceed the size of the pore, and thus its shape and monomer density are not heavily distorted by the channel, $h(w)$ obtained in this way may serve as a good approximation.

An alternative way to determine $\langle N(w) \rangle$, and thus $h(w)$, is based on the observation that the reduction in solution conductivity is proportional only to the wt% concentration of polymer, that is, to monomer density, and does not depend on polymer M_r (ref. 5). Supposing that the relative reduction in the pore conductance is also proportional to the density of monomers inside the pore, we calculate pore-bulk polymer partitioning as

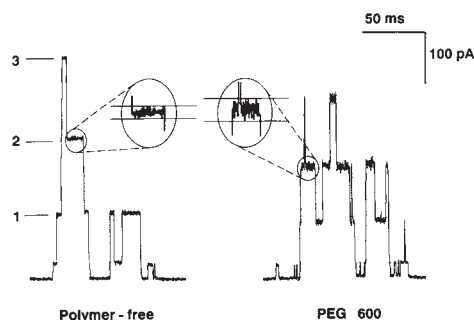


FIG. 2 Ion current of a single alamethicin channel as changed by polymer addition. (Left, polymer-free; right, after addition of PEG 600). PEG 600 decreases the amplitude of different conductance levels (numbers 1, 2 and 3 at left) and induces additional current noise. Current records filtered by an 8-pole Bessel filter at 3 kHz are shown enlarged (2 $\times$) in insets for comparison. Different conductance levels of the channel correspond to different numbers of closely packed molecular pores (Fig. 1) of nearly uniform dimensions^{5,20}. The time/current scale in the right upper corner is common to both records.

a function of polymer M_r to obtain $\langle N(w) \rangle$. Both methods give similar results; we use this second method as it employs less restrictive assumptions. It is also reassuring that polymer partitioning into the pore can be verified independently by measuring size-modulated osmotic response⁸.

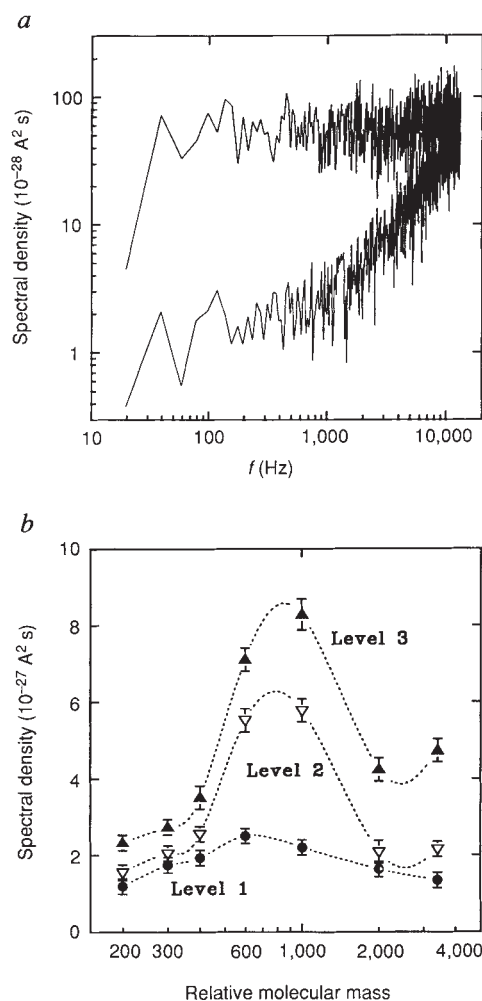


FIG. 3 Measurement of polymer-induced current fluctuations. *a*, Spectrum of open-channel noise in the presence of PEG 600 at level 2 (upper trace) is shown in comparison with the spectrum of the background noise obtained from the parts of the current recordings when the channel was closed (lower trace). *b*, Amplitude of the 'white' spectral part of the open-channel noise changes with polymer relative molecular mass, M_r . The spectrum was averaged over a 200–2,000 Hz frequency range with the background subtracted.

METHODS. In spectral measurements an 8-pole Butterworth filter was employed, the corner frequency being chosen to be three-eighths of the sampling frequency (40 kHz). To separate open-channel noise from a strong low-frequency contribution of the channel switching between different conductance states, we first recorded the segments of the current corresponding to a particular conductance state into the computer memory; then, after cutting 0.25 ms off both ends of each segment to eliminate transient currents and after zeroing the mean value of each segment, we connected them into a 2,048-point vector to be used in Fourier transformation. This signal processing procedure is justified by the fact that characteristic times of current fluctuations within a particular state are orders of magnitude smaller than inter-state switching times. Nevertheless, to check for possible spectrum distortions due to a limited lifetime of the channel at a particular state, we performed a specially designed test. A signal from a calibrated noise generator was admixed to the channel current recording and processed as described above. The spectra so obtained were then compared to those measured directly from the generator output. It turned out that the damping of the admixed noise spectra is significant only at frequencies lower than 50 Hz.

The mean value of square conductance fluctuations, $\langle h(w)^2 \langle N(w) \rangle \rangle$, can also be calculated theoretically as a product of the low-frequency limit of conductance fluctuation spectral density $S_h(0)$ times the spectral width created by diffusion relaxation times. General considerations show that the bandwidth of these fluctuations should be proportional to the diffusion coefficient D of polymer inside the pore. To calculate the bandwidth for the case of a long pore we used a one-dimensional approximation. For the probability $P(x_0, x, \tau)$ to find a particle in position x at time τ if it had been in position x_0 at $\tau=0$, we have

$$D(\partial^2 P(x_0, x, \tau)/\partial x^2) = \partial P(x_0, x, \tau)/\partial \tau$$

For a narrow channel, the probability that a particle that has left the channel will return diminishes rapidly as the particle moves from the mouth of the channel. Placing absorbing boundaries at both ends of the capillary, that is at $x=0$ and $x=L$, we solve the problem using methods described elsewhere¹⁰.

Formally, for the spectral density, $S_h(0)$, of conductance fluctuations at frequency $f=0$, we obtain

$$S_h(0) = S_h(f)|_{f=0} = \langle (\delta h)^2 \rangle L^2 / 3D$$

where $\langle (\delta h)^2 \rangle$ is the time-averaged square fluctuation. The diffusion-generated bandwidth of pore conductance fluctuations is $3D/L^2$. The corresponding time constant that describes averaged relaxation of polymer number fluctuation inside the pore is equal to $L^2/6\pi D$. This value is remarkably close to the intuitive result $L^2/12D$ reported by Feher and Weissman¹¹.

Experimentally, we obtain $S_h(0) = S_I(0)/V^2$ from the measured low-frequency 'white' part of the current noise spectrum $S_I(0)$ and transmembrane voltage V . Equating time-averaged and ensemble-averaged square fluctuations, $\langle (\delta h)^2 \rangle$ and

$h(w)^2 \langle N(w) \rangle$, for the polymer diffusion coefficient inside the pore we have

$$D(w) = h(w)^2 L^2 \langle N(w) \rangle / 3S_I(0)$$

The diffusion coefficients calculated for polymers of different M_r are shown in Fig. 4, together with their values in bulk aqueous solution. We assume channel length $L=40$ Å and radius $R=6.5$ Å, based on results of measurements⁵ and on molecular model considerations⁶.

The smallest polymer used in our experiments, PEG 200, has a gyration radius of 4.0 Å (ref. 12), already quite close to the pore radius. Formal application of restricted-diffusion theory¹³ in this case gives the right order of magnitude for the reduction in diffusion rate (Fig. 4) even though its application is highly questionable here. Restricted-diffusion theory is developed for hard spheres, whereas PEG polymers in aqueous solutions exist as well-characterized random coils¹⁴. The observed order-of-magnitude-slower polymer diffusion in the pore can have many causes: size restriction, friction with the walls, higher viscosity of water in the pore, and so on. The absence of any significant dependence of the diffusion coefficient on the polymer M_r disagrees with the expected results of restricted-diffusion considerations for hard spheres, and might reflect the confinement of flexible coil in a pore whose dimensions are comparable to the size of the polymer. Although the viscosity of water inside the channel may be different from the bulk, the observed decrease in diffusion coefficient does not provide a quantitative measure of this.

It is important to note here that the diffusion coefficient obtained in our experiments describes transport of polymers only once they are inside the pore. To describe the overall diffusion of polymers across the whole membrane, one should take into account partitioning of polymers between the pore and bulk solution. Partitioning of polymers in the pore quickly decreases with increasing polymer M_r (Fig. 4); so does diffusion across the membrane.

Fluctuation analysis of currents through ion channels, performed to obtain dwell times of charged permeant species, has been successfully used previously. Classic examples include studies on Ca^{2+} block of monovalent currents through calcium channels¹⁵ and on Na^+ block of proton currents through the gramicidin A channels¹⁶. Now that strong evidence for the existence of protein-conducting channels is starting to emerge¹⁷, the importance of studies on neutral-permeant transport is obvious. Using the approach outlined here, it should be possible to conduct such studies at a single-channel level. □

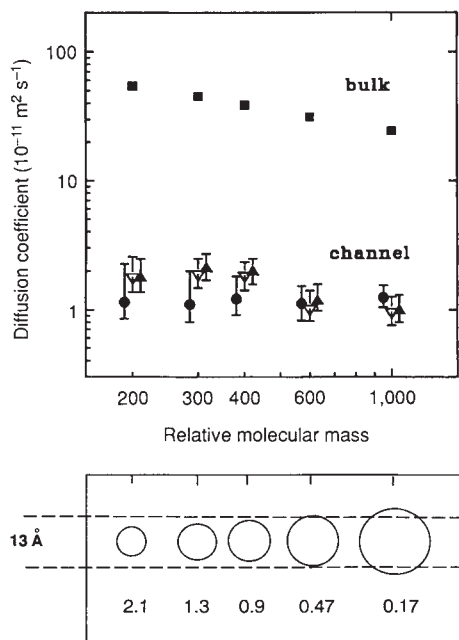


FIG. 4 Top, Diffusion coefficients of polymers inside the channel calculated from polymer-induced current noise (symbols for different levels correspond to those in Fig. 3b). An order of magnitude decrease in comparison with their values in bulk water is already expected from restricted-diffusion theory¹³ for the smallest polymer, PEG 200. Bottom, PEG gyration diameters are compared with the size of the alamethicin channel opening (13 Å). Numbers below the gyration diameters correspond to the average number of PEG molecules in the channel (level 1).

Received 8 March; accepted 6 June 1994.

- Kubitschek, H. E. *Nature* **182**, 234–235 (1958).
- Allen, T. in *Particle Size Analysis* 110–127 (Society for Analytical Chemistry Publishers, London, 1967).
- Bunville, L. G. in *Modern Methods of Particle Size Analysis* (ed. Barth, H. G.) 1–42 (Wiley, New York, 1984).
- DeBlois, R. W., Bean, C. P. & Wesley, R. K. A. *J. Colloid Interface Sci.* **61**, 323–335 (1977).
- Bezrukov, S. M. & Vodyanov, I. *Biophys. J.* **64**, 16–25 (1993).
- Sansom, M. S. P. *Eur. Biophys. J.* **22**, 105–124 (1993).
- Hall, J. E., Vodyanov, I., Balasubramanian, T. M. & Marshall, G. R. *Biophys. J.* **45**, 233–247 (1984).
- Vodyanov, I., Bezrukov, S. M. & Parsegian, V. A. *Biophys. J.* **65**, 2097–2105 (1993).
- Krasilnikov, O. V., Sabirov, R. Z., Ternovsky, V. I., Merzliak, P. G. & Muratkodjaev, J. N. *FEMS Microbiol. Immunol.* **105**, 93–100 (1992).
- Bezrukov, S. M. & Vodyanov, I. in *Membrane Electrochemistry* (Advances in Chemistry Ser. No. 235, American Chemical Soc., Washington DC, in the press).
- Feher, G. & Weissman, M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 870–875 (1973).
- Kuga, S. *J. Chromatogr.* **206**, 449–461 (1981).
- Bean, C. P. in *Membranes* (ed. Eisenman, G.) 1–54 (Dekker, New York, 1972).
- Couper, A. & Stepto, R. F. T. *Trans. Faraday Soc.* **65**, 2486–2496 (1969).
- Hess, P. & Tsien, R. W. *Nature* **309**, 453–456 (1984).
- Heinemann, S. H. & Sigworth, F. J. *Biochim. biophys. Acta* **987**, 8–14 (1989).
- Simon, S. M. & Blobel, G. *Cell* **65**, 371–380 (1991).
- Montal, M. & Mueller, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3561–3566 (1972).
- Balasubramanian, T. M. et al. *J. Am. chem. Soc.* **103**, 6127–6132 (1981).
- Gordon, L. G. M. & Haydon, D. A. *Phil. Trans. R. Soc. B* **270**, 433–447 (1975).

ACKNOWLEDGEMENTS. This work was supported by the US Office of Naval Research.

Effect of increasing atmospheric methane concentration on ammonium inhibition of soil methane consumption

Gary M. King & Sylvia Schnell

Darling Marine Center, University of Maine, Walpole, Maine 04573, USA

SOILS currently consume about 30–40 Tg methane per year^{1,2}, which is comparable to the net annual increase in atmospheric methane concentration from 1980 to 1990³. Most soils consume methane^{2,4–9}, but the extent varies with soil water content, land use and ammonium inputs^{5,10–13}. Ammonium concentrations in many soils have increased in recent years as a result of land-use changes and increases in ammonium concentration in precipitation^{14,15}. Ammonium strongly inhibits soil methane consumption, but the mechanism is uncertain. Even if enhanced ammonium concentrations are subsequently reduced, inhibition can still persist for months to years^{12,13}. Here we show, from field and laboratory experiments, that the extent of ammonium inhibition increases with increasing methane concentration. We propose that nitrite formation from methanotrophic ammonium oxidation accounts for much of the observed inhibition, and that the persistence of inhibition with reduced ammonium concentrations is due to the limited capacity of methanotrophs to grow or recover in present concentrations of atmospheric methane. We suggest that past increases in atmospheric methane concentration may have increased the inhibitory effect of ammonium, thereby decreasing soil methane uptake capacity, and that this mechanism could also provide a positive feedback on future atmospheric methane concentrations.

The methane consumption of *in situ* forest soil was significantly inhibited by fertilization with ammonium (Fig. 1). Inhibition occurred within 24 h, and persisted unabated during a 12-d assay period. Rapid and persistent (months-years) effects of ammonium additions on methane consumption, as well as on microbial biomass and respiration more generally, have been reported previously for a variety of soils^{5,12,13,16}. For example, Nesbit and Breitenbeck¹³ reported inhibition in soils for up to 30 d, even though >75% of the added ammonium had been nitrified within 5 d; Mosier *et al.* observed inhibition in fertilized grasslands over periods from ~1 to 9 yr after fertilization¹². Results of an analysis of methane consumption by forest soils after clear-cutting and transformation to pasture use suggest that diminished uptake can persist for even longer periods¹⁷.

Results from our field and laboratory analyses indicate that the inhibitory effects of ammonium depend on the concentrations of methane oxidized. In our field experiments, the extent of ammonium inhibition was directly related to the methane concentrations maintained in the headspace of chambers deployed in fertilized and unfertilized plots (Fig. 1). Inhibition by ammonium was least with 0.9 p.p.m., and greatest with 5 p.p.m. methane concentrations. Similar results have been obtained for soils incubated in the laboratory: inhibition of methane uptake by ammonium increased from 44.5% to 72.0% with methane headspace concentrations from 0.6 to 10 p.p.m. (Fig. 2). The greater absolute magnitude of inhibition in the laboratory analyses probably reflects the use of small soil samples (10 g) to which ammonium was added directly; this treatment probably resulted in higher local concentrations of ammonium than in the field study, where ammonium was applied at the soil surface. Overall, the pattern for changes in

inhibition as a function of methane concentration is comparable for soils incubated *in vitro* or *in situ*.

These results can be attributed to the physiology of methane oxidation. Methanotrophs oxidize both methane and ammonia, forming methanol or hydroxylamine, respectively, using one of two methane monooxygenases (MMO)¹⁸. As shown below, these reactions require a source of reducing equivalents (NADH + H⁺ or reduced pyrrolo-quinoline quinone)¹⁸. Methanol oxidation, for example, produces more reductant per mole than is initially required per mole of methane. Although hydroxylamine (a potential inhibitor of MMO¹⁹) is further metabolized to nitrite

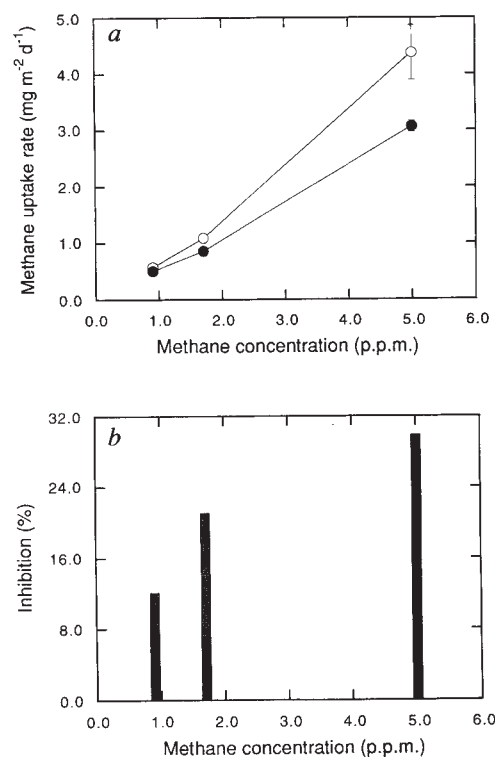
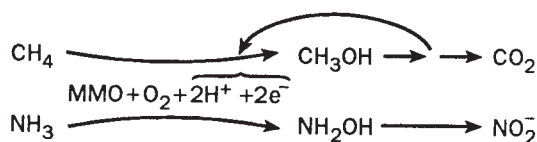


FIG. 1 Effect of methane concentration on ammonium inhibition of methane oxidation. Two adjacent 3 × 1 m plots were established in a mixed oak–pine forest near the Darling Marine Center, Walpole, Maine. Clayey-loam soils were overlain by an organic horizon and litter layer ~7 cm thick. One plot was fertilized on 29 September 1993 by spraying it evenly with 6.25 l of aqueous 11.4 mM ammonium chloride. This was equivalent to 6 mm rainfall, and yielded a nitrogen application of 71 mmol m⁻², a level comparable to atmospheric deposition in areas subjected to agriculturally-derived pollution²⁷ and to the higher fertilization treatment of Steudler *et al.*⁵ The second plot was treated similarly with deionized water as a control. Within 2–3 h after spraying, 11.5 cm inner diameter polyvinyl chloride chambers were deployed in each plot. The chambers were sealed with end-caps modified with inlet, outlet and sampling ports and inserted to a depth of 17 cm, leaving a headspace of ~1 l. The chambers were connected to one of three air–methane sources with flow rates of 300 ml min⁻¹ per chamber. The air–methane sources were regulated to maintain methane concentrations in the chamber headspaces of 0.9, 1.7 or 5.0 p.p.m. during the following 9 d. During analysis of methane uptake rates, flows for all chambers were stopped, inlet and outlet ports were sealed, and samples were removed at intervals for ~6 h (by needle and syringe) for analysis of methane content as described previously²⁸. Soil temperatures were ~11 °C during the analysis period. a, Means of triplicate (5 p.p.m.) or quadruplicate (0.9 and 1.7 p.p.m.) determinations of methane uptake rates for control (empty circles) and ammonium-fertilized (filled circles) plots sampled on day 4. Control and fertilized plot uptake rates were significantly different ($P < 0.05$) according to a Mann–Whitney U test. Error bars are ± 1 standard error. b, Inhibition of methane uptake for each headspace methane concentration, calculated from the ratio between control and fertilized plots.

by methanotrophs^{18,20,21}, recycling of reductant to MMO is apparently inefficient relative to recycling



from methanol²². Consequently, ammonia oxidation can become reductant-limited; this accounts for the stimulation of ammonia oxidation by additions of suitable co-substrates, for example, formate^{18,20}. Our results also show that increasing the concentration of methane from 1.7 to 500 p.p.m. enhances nitrite production in cultures of *Methylomonas albus*; the inhibitory effect of ammonium on this culture increases from 1.7–100 p.p.m. methane, and is reversed somewhat at higher methane concentrations (Fig. 3). As the kinetics of MMO favour oxidation of methane over ammonia¹⁸, increased ammonium inhibition with increasing methane probably results from the inhibitory effect of nitrite, the final product of ammonia oxidation by methanotrophs. Both soil and culture data (Fig. 4) show that nitrite strongly inhibits methane consumption. The mechanism for the effect of nitrite is not clear, but it can inhibit formate dehydrogenase²³, an important NADH + H⁺-producing enzyme in the catabolism of methane¹⁸. Inhibition by nitrite is also substrate-dependent, in that a strong response is observed at low methane levels (≤ 100 p.p.m.), with decreasing inhibition at

higher concentrations (>100 p.p.m.). This latter observation probably explains the decreased sensitivity to ammonia of methane oxidation in cultures at elevated methane concentrations (Fig. 3).

Collectively, these results provide a model for understanding the effects of ammonium on soil methane consumption, and for predicting responses to ammonium fertilization. The mechanism of inhibition by ammonium probably involves competition between methane and ammonia for MMO, as has been described for pure cultures^{18,20,21}. But nitrite causes an additional, non-competitive inhibition that seems to decrease the cell-specific activity or biomass of soil methanotrophs. The diminished capacity for methane consumption due to ammonia fertilization persists for short (days) and long (months–years) periods, even after ammonium concentrations are reduced to background levels^{5,12,13}. This is perhaps attributable to the limited capacity of soil methanotrophs to grow *in situ* (S.S. and G.M.K., manuscript in preparation).

In order to assess the growth potential of soil methanotrophs, we exposed the most active methane-consuming layer (4–8 cm depth) of a forest soil to methane concentrations significantly above those found within the layer *in situ* (~ 1 p.p.m.). Soils were incubated at ambient laboratory temperatures in a chamber that was continuously flushed with air containing constant concentrations of water vapour and methane. Flow rates were adjusted to prevent methane depletion within the chamber. Rates of atmospheric methane uptake were assayed periodically by measuring methane depletion by soils incubated with static headspaces; soils from flow-through chambers with methane concentrations >1.7 p.p.m. were equilibrated with ambient air before the assay. The capacity of the soils to consume methane did not increase during relatively long-term incubations (weeks–months), even though absolute uptake rates increased with increasing methane concentration. For example, methane uptake rates for soils maintained with 1.7 p.p.m. methane did not vary significantly during a 126-d period (mean for nine time points, 4.9 ± 0.3 mg m⁻² d⁻¹); similarly methane uptake rates for soils

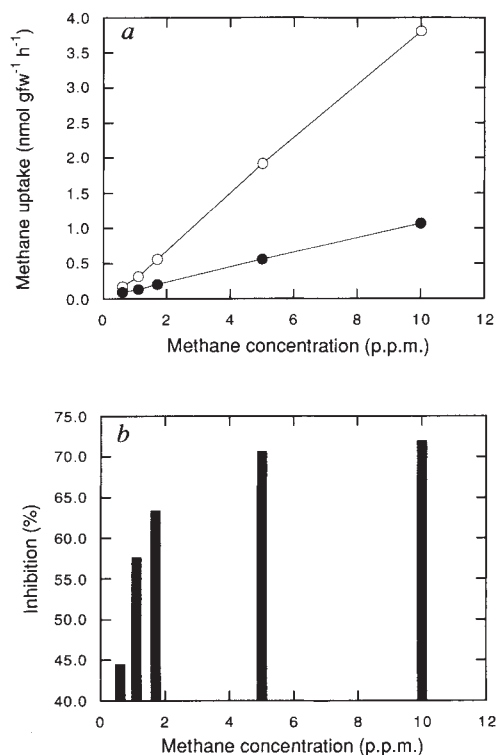


FIG. 2 Inhibition by ammonium chloride of methane consumption in forest soil with various headspace methane concentrations. Fresh soils (10 g; from 7–10 cm depth) were incubated in stopper-sealed 120-ml glass bottles. Quadruplicate samples for each methane concentration were treated with either 1 ml of deionized water or an aqueous solution of ammonium chloride yielding a final concentration of 1 μ mol per gram fresh weight (gfw). Methane uptake was followed over time by removing headspace samples at intervals with a needle and syringe. a, Uptake rates for each of the paired controls (empty circles) and ammonium-treated soils (filled circles) were significantly different ($P < 0.02$; Mann–Whitney U-test), and coefficients of variation for each set of replicates were all $<2\%$. b, Inhibition for each methane concentration, calculated from the ratio between control and fertilized soils.

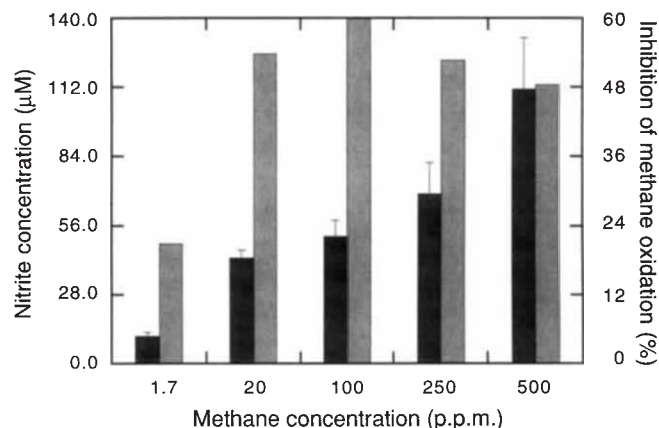


FIG. 3 Inhibition of methane oxidation by ammonium (light grey), and nitrite production (dark grey) at various methane concentrations in cultures of the Type I methanotroph *Methylomonas albus*. A methane-grown culture was collected during logarithmic growth by centrifugation. Cells were resuspended (absorbance at 600 nm = 0.3) in a phosphate-buffered mineral-salts medium²⁸ with or without 500 μ M ammonium chloride; 10-ml volumes were incubated in 160-ml sealed serum bottles that contained various initial methane concentrations. The inhibitory effect of ammonium was calculated from estimates of methane oxidation at each of the initial methane concentrations in the presence and absence of ammonium; oxidation rates were determined by removing headspace samples at intervals for analysis by gas chromatography. Nitrite concentrations were determined colorimetrically using subsamples collected at the end of the incubation. All assays were performed in triplicate; error bars for nitrite concentrations are ± 1 standard error.

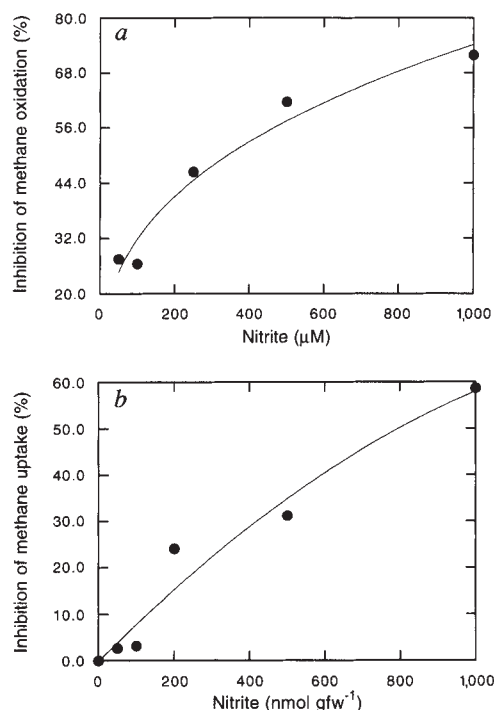


FIG. 4 *a*, Inhibition of methane oxidation by nitrite in cultures of *M. albus*. Samples (10 ml) of cultures were prepared as described in Fig. 3 legend, except that nitrite concentrations were varied from 50 to 1,000 μM and the initial headspace methane concentrations were 100 p.p.m. Initial methane uptake rates in the absence of added nitrite were 11.8 ± 1.1 nmol per ml culture h^{-1} . *b*, Inhibition of methane uptake by forest soils prepared as described in Fig. 2 legend, except that the soils were amended with an aqueous solution of sodium nitrite to give a range of final concentrations from 0 to 1,000 nmol per gram fresh weight (gfw) soil. Uptake rates were assayed by following the depletion of atmospheric methane, and inhibition was expressed relative to an unamended control.

maintained with 17 p.p.m. methane did not vary during a 14-d incubation (mean for three time points, 27.6 ± 2.4 $\text{mg m}^{-2} \text{d}^{-1}$). The constancy of uptake in both cases suggested that populations of soil methanotrophs could not increase even with an increase in methane availability up to 17 times greater than that *in situ*. However, the capacity of consumption approximately doubled during incubations of 18 d with 1,000 p.p.m. methane. This response indicated that growth on methane was not limited by nutrients or other factors in the soil. These observations are consistent with results of Bender and Conrad²⁴, who were only able to enrich and isolate soil methanotrophs by using elevated methane concentrations.

We also propose that ammonium inhibition accounts for the sub-surface maxima of atmospheric methane consumption in soils^{6,8,25}. The low activities in surface soils are enigmatic, because simple growth models for limiting, diffusible substrates (see, for example Koch²⁶) predict maximal activities near the soil-air interface, where methane concentrations are highest. Sub-surface maxima could be a consequence of inhibition in the uppermost soil horizons, where both deposition and ammonium regeneration are typically highest. An exception to the typical pattern for the depth distribution of methane consumption has been reported by Whalen and Reeburg⁷, who found highest activities near the surface of tundra soils. This pattern may be attributable to lower rates of ammonium deposition or regeneration in tundra compared to temperate forest or grassland soils.

The interactions between atmospheric methane concentrations and ammonium inhibition described here can constrain the magnitude of the soil methane sink. At present, soil methane consumption behaves as a first-order process; rates of consumption are determined by atmospheric methane partial pressures at the soil surface, and apparent rate uptake constants that integrate the depth distribution and activity of methanotrophs, seasonality and other factors specific for given systems (for example, $dM/dt = -kM$, where t is time, k is an apparent uptake rate constant with units of $\text{m}^{-2} \text{time}^{-1}$ and M is a methane concentration). Our results indicate that elevated methane partial pressures increase the extent of ammonium inhibition, and decrease apparent uptake rate constants. The relationship between apparent uptake rate constants and atmospheric methane concentrations in the presence of ammonium can thus be described as $dk/dM < 0$; in the absence of ammonium, the

relationship conforms to $dk/dM \geq 0$. As a consequence, soil methane uptake rates increase with elevated atmospheric methane concentrations, but the increases are less in the presence of inhibitory levels of ammonium than in its absence.

Because ammonium inputs to soils have been increasing²⁷ along with increases in atmospheric methane concentrations, we suggest that the capacity for soil methane consumption has decreased during the post-industrial period, and that it could continue to do so in the future. Models for global methane budgets must therefore consider changes in regional water balances, land use, soil nitrogen dynamics and the positive feedback between methane and ammonium inhibition to determine accurately the contribution of soils to atmospheric methane levels. □

Received 6 December 1993; accepted 10 June 1994.

- Dörr, H., Katruff, L. & Levin, I. *Chemosphere* **26**, 697–705 (1993).
- Striegl, R. G., McConnaughey, T. A., Thorstenson, D. C. & Woodward, J. C. *Nature* **357**, 145–147 (1992).
- Steele, L. P., et al. *Nature* **358**, 313–316 (1992).
- Keller, M., Groeau, T. J., Wofsy, S. C., Kaplan, W. A. & McElroy, M. B. *Geophys. Res. Lett.* **10**, 1156–1159 (1983).
- Stedler, P. A., Bowden, R. D., Mellilo, J. M. & Aber, J. D. *Nature* **341**, 314–316 (1989).
- Born, M., Dörr, H. & Levin, I. *Tellus* **42B**, 2–8 (1990).
- Whalen, S. C. & Reeburg, W. S. *Nature* **346**, 160–162 (1990).
- Adamsen, A. P. S. & King, G. M. *Appl. envir. Microbiol.* **59**, 485–490 (1992).
- Crill, P. M. *Globl. Biogeochem. Cycles* **4**, 319–334 (1991).
- Whalen, S. C., Reeburg, W. S. & Sandbeck, K. A. *Appl. envir. Microbiol.* **56**, 3405–3411 (1990).
- Keller, M., Mitre, M. E. & Stallard, R. F. *Globl. Biogeochem. Cycles* **4**, 21–27 (1990).
- Mosier, A., Schimel, D., Valentine, D., Bronson, K. & Parton, W. *Nature* **350**, 330–332 (1991).
- Nesbit, S. P. & Breitenbeck, G. A. *Agric. Ecosys. Envir.* **41**, 39–54 (1992).
- Nihlgård, B. *Ambio* **14**, 2–8 (1985).
- Willey, J. D. & Kiefer, R. H. *J. Elisha Mitchell Sci. Soc.* **109**, 1–19 (1993).
- Söderström, B., Bååth, E., & Lundgren, B. *Can. J. Microbiol.* **29**, 1500–1506 (1983).
- Keller, M., Veldkamp, E., Weitz, A. M. & Reiners, W. A. *Nature* **365**, 244–246 (1993).
- Bédard, C. & Knowles, R. *Microbiol. Rev.* **53**, 68–84 (1989).
- Hubley, J. H., Thomson, A. W. & Wilkinson, J. F. *Arch. Mikrobiol.* **95**, 365–368 (1975).
- O'Neill, J. G. & Wilkinson, J. F. *J. gen. Microbiol.* **100**, 407–412 (1977).
- Dalton, H. *Arch. Mikrobiol.* **114**, 273–279 (1977).
- Yoshinari, T. *Can. J. Microbiol.* **31**, 139–144 (1985).
- Jollie, D. R. & Lipscomb, J. D. *J. biol. Chem.* **266**, 21853–21863 (1991).
- Bender, M. & Conrad, R. *FEMS Microbiol. Ecol.* **101**, 261–270 (1992).
- Koschorreck, M. & Conrad, R. *Globl. Biogeochem. Cycles* **7**, 109–121 (1993).
- Koch, A. L. *Adv. microb. Ecol.* **11**, 37–70 (1990).
- Georgii, H. W., Resseke, C. & Rohbeck, E. *Atmos. Envir.* **18**, 581–589 (1984).
- King, G. M. & Adamsen, A. P. S. *Appl. envir. Microbiol.* **58**, 2758–2763 (1992).

ACKNOWLEDGMENTS. This work was supported by the US NSF, NASA and a Deutsche Forschungsgemeinschaft fellowship (S.S.).

Evidence for metastable equilibrium between hydrocarbons under hydrothermal conditions

Jeffrey S. Seewald

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

THERMAL maturation of organic matter in sedimentary basins results in the generation of numerous organic alteration products. These products influence the chemical and physical alteration processes of sediments¹⁻³, and in some instances their accumulation results in the formation of oil and natural-gas deposits^{4,5}. Much uncertainty exists about the factors that control the relative abundances of individual organic species during maturation. Although it is clear that kinetic barriers allow thermodynamically unstable species to persist in a metastable state for geologically significant periods of time, local equilibrium between more-reactive species may substantially influence their abundance. Here I report the results of redox-buffered hydrothermal experiments designed to investigate reactions that may occur in geological environments between ethane, ethene, water and inorganic redox-sensitive minerals. I find that reversible metastable thermodynamic equilibrium is attained between these species. Because water participates directly in this equilibrium, it may represent a reactive and abundant source of hydrogen for hydrocarbon generation in sedimentary basins. This demonstration that metastable equilibrium is attained implies that some organic geochemical reactions can be thermodynamically modelled and predicted despite the absence of total thermodynamic equilibrium.

Oil and natural-gas generation are generally viewed as kinetically controlled processes in which the absolute abundance and relative distribution of organic species reflect the relative rates of production and destruction and/or the primary composition of precursor material^{4,5}. The possibility, however, that reactions characterized by low activation energies attain a state of thermodynamic equilibrium while other reactions, with relatively high activation energies, remain out of equilibrium has important implications for understanding chemical processes during sediment alteration. For example, because reactions controlled by thermodynamic equilibrium are affected solely by energy changes, they are independent of reaction mechanism and time and respond only to temperature, pressure and composition of the chemical system. Thus, organic compounds that attain a state of thermodynamic equilibrium can be used to constrain chemical and physical conditions. Such an approach has

been employed to define geochemical processes relevant to the stability of organic species in basinal brines and petroleum^{6,7}, amino acids in meteorites⁸ and the origin of life in hydrothermal systems⁹. These studies, however, are based on theoretical calculations and have never been tested experimentally.

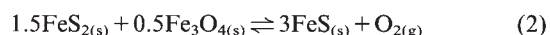
The experiments presented here were conducted at 325 and 350 °C at 350 bar. Although a wide variety of subsurface temperatures exist in sedimentary basins, those associated with extensive oil and natural-gas generation are generally <350 °C. For example, in sediment-covered ridge crest hydrothermal systems such as Guaymas Basin, Escanaba Trough, and Middle Valley, oil and natural-gas generation is a common phenomenon in the temperature range 200–350 °C (refs 10–14). Oil and natural-gas in economic deposits are generally believed to form at 50–150 °C and 150–220 °C, respectively^{4,5,15}. However, recent investigations of deep oil and natural-gas deposits indicate that temperatures approaching 300 °C may be more common than previously believed¹⁶.

The primary objective of this study was to constrain the extent to which thermodynamic equilibrium may regulate the relative abundance of *n*-alkanes and *n*-alkenes at elevated temperatures. Specifically, equilibration of ethene (C₂H₄) and ethane (C₂H₆) according to the reaction:



was investigated. The *n*-alkanes represent an important class of compounds constituting oil and natural gas. Numerous hydrous and anhydrous pyrolysis experiments have demonstrated the generation of low-molecular-weight alkenes as unstable reaction intermediaries in addition to abundant alkanes during the thermal maturation of kerogen^{17–21}. Because hydrogenation of alkenes may represent an important source of alkanes, understanding factors controlling such reactions and identifying sources of hydrogen are crucial to models of hydrocarbon generation. Although there are some notable exceptions^{6,21,22}, kerogen is generally considered to be the only source of hydrogen for hydrocarbon formation⁵.

Owing to the different oxidation states of carbon in ethane and ethene, it is essential to constrain redox conditions for an unambiguous interpretation of the factors controlling their stability. A pyrite–pyrrhotite–magnetite (PPM) mineral assemblage was employed to buffer oxygen fugacity according to the reaction:



where FeS₂, Fe₃O₄ and FeS are pyrite, magnetite and pyrrhotite, respectively. In an aqueous environment, the activity of H_{2(aq)} is also buffered by the PPM assemblage due to the disproportionation of water as follows:

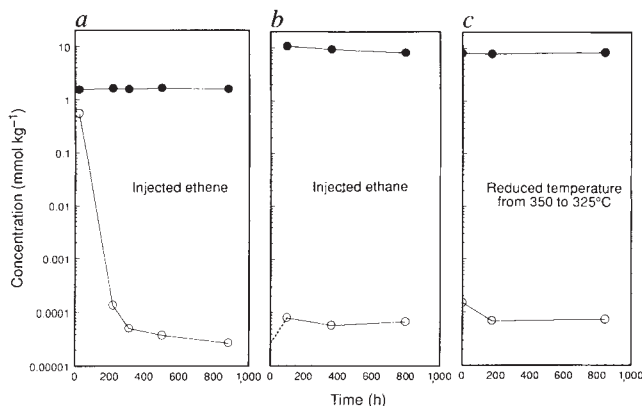


FIG. 1 Variations in the dissolved concentrations of ethane (filled circles) and ethene (open circles) with time during reaction with a pyrite–pyrrhotite–magnetite mineral redox buffer at 325 °C and 350 bar after perturbing the chemical or physical conditions of the experimental system. (Details of perturbations shown on Figure.)

TABLE 1 Dissolved concentrations of selected aqueous species

Time (h)	T (°C)	C ₂ H ₄	C ₂ H ₆	CO ₂	CH ₄	H ₂	H ₂ S	log[C ₂ H ₆ /C ₂ H ₄]
(All concentrations mmol per kg-fluid)								
Experiment 1: C ₂ H ₄ and H ₂ O injected								
0*	325	≥0.55	≤2.7	≤0.087	≤0.04	≤0.1	≤1.8	
24.4	325	0.55	1.5	0.14	0.083	0.31	4.8	0.44
217.7	325	0.00014	1.6	0.25	0.17	0.36	7.3	4.06
313.2	325	0.000051	1.6	0.28	0.18	—	—	4.50
502.2	326	0.000038	1.6	0.35	0.23	0.35	8.3	4.62
885.7	325	0.000027	1.6	0.54	0.31	0.36	6.1	4.77
Experiment 2: C ₂ H ₆ and H ₂ O injected								
0*	325	<0.000027	≥11	≤0.22	≤0.11	≤0.12	≤1.8	≥5.6
100.6	325	0.000080	11	0.19	0.11	0.29	3.1	5.14
363.3	326	0.000059	9.3	0.23	0.11	0.31	3.3	5.20
797.6	326	0.000068	8.0	0.26	0.12	0.31	4.4	5.07
Experiment 3: Temperature increased from 325 to 350 °C and H ₂ O injected†								
0*	350	<0.000015	≤0.28	≤0.50	≤0.16	≤0.16	≤2.2	≥4.4
287.8	350	0.000029	0.21	0.52	0.16	0.47	8.2	3.86
622.0	351	0.000022	0.21	0.73	0.22	0.51	9.3	3.99
Experiment 4: C ₂ H ₆ and H ₂ O injected								
0*	351	<0.0000089	≥8.8	≤0.18	≤0.031	≤0.23	≤3.5	≥6.0
23.5	350	0.00028	8.8	0.18	0.066	0.35	6.0	4.50
55.8	351	0.00020	8.5	0.17	0.066	—	—	4.63
839.0	350	0.00015	8.0	0.56	0.26	0.43	10	4.73
Experiment 5: Temperature reduced from 350 to 325 °C								
0	325	0.00015	8.0	0.56	0.26	0.43	10	4.73
176.5	325	0.000069	7.7	0.55	0.28	—	—	5.05
847.3	326	0.000074	8.3	0.60	0.31	0.34	4.9	5.05

All the above dissolved concentrations of selected aqueous species were determined during reaction with a pyrite–pyrrhotite–magnetite (PPM) mineral redox buffer at 350 bar. Gas chromatography was used for all determinations, except for H₂S which was determined gravimetrically by precipitation as Ag₂S. Analytical uncertainty is ±5% for all species.

* Fluid compositions were analyzed before H₂O and gas injection. Concentrations at 0 h were corrected for dilution based on the amount of H₂O injected. Due to uncertainties in the mass of H₂O in the reaction cell at the time of injection, however, the concentrations are absolute maximums. Technical limitations preclude accurate determination of the amount of gas injected. It has been assumed that the concentration of injected gas at 0 h is ≥ the concentration in the first sample.

† Thermocatalytic degradation of minor reduced carbon contaminants present in the PPM mineral assemblage is partially responsible for the increases in dissolved CO₂ and CH₄ concentrations after increasing temperature to 350 °C. Before starting experiments 1, 2 and 3 at 325 °C, the mineral assemblage was allowed to react until CO₂ and CH₄ production from this source ceased.

According to reaction (1), the relative abundance of ethane and ethene depends directly on the activity of H_{2(aq)}. Thus under equilibrium conditions, the ethane/ethene activity ratio is fixed at a unique value at a given temperature and pressure by the PPM redox buffer.

The experiments were conducted in a flexible gold reaction cell hydrothermal apparatus²³. This equipment allows fluid samples to be added to or removed from the reaction cell during an experiment without disturbing the temperature and pressure condition. After equilibration of the PPM mineral assemblage at 325 °C, individual experiments were initiated by injecting either ethane or ethene into the reaction cell along with degassed water. Ethane–ethene equilibrium (reaction (1)) was then monitored by measuring dissolved gas concentrations as a function of time (Table 1, Fig. 1). In experiments 3 and 5 (Table 1) the effects of temperature on ethane and ethene stability were examined. The reaction cell was not dismantled between experiments. Consequently, the concentrations of dissolved gases at 0 h reflect the reaction cell contents at the termination of the previous experiment in addition to any added reactants.

Dissolved H₂ concentrations at 325 °C repeatedly and rapidly returned to a value of 0.34 ± 0.03 mmol per kg-fluid after injection of water and ethane or ethene into the reaction cell (Table 1). This concentration is consistent with values predicted for reactions (2) and (3) at thermodynamic equilibrium^{24–26} and demonstrates that the disproportionation of water is a relatively rapid process. Thus, water represents an abundant source of hydrogen that is available for reaction with organic species at elevated temperatures and pressures.

Following injection of ethene at 325 °C, the dissolved concentration of this gas decreased by approximately four orders of

magnitude before asymptotically approaching a nearly constant value (Table 1, Fig. 1a). In contrast, dissolved ethene concentrations increased by at least a factor of two following injection of ethane (Table 1, Fig. 1b), consistent with formation of ethene according to reaction (1). Examination of Fig. 2 reveals that, the large variations in absolute concentrations notwithstanding, the ethane/ethene concentration ratio approached a constant value ($\log[\text{ethane/ethene}] = 4.95 \pm 0.15$) after being externally adjusted to higher and lower values by injection of ethane and ethene, respectively. These data provide strong evidence that the relative abundance of ethane and ethene are regulated by reversible thermodynamic equilibrium. The observed ethane/ethene concentration ratio is consistent with predicted values based on theoretical thermodynamic data for reactions (1), (2) and (3) at 325 °C and 350 bar (refs 24–27) (Fig. 2), the 0.61 log unit difference being within the uncertainty associated with the theoretical predictions.

Increasing the temperature from 325 to 350 °C resulted in an initial increase in the concentration of dissolved ethene, which caused a decrease in the ethane/ethene concentration ratio (Table 1). Increased ethene concentration suggests greater stability relative to ethane with increasing temperature in the presence of PPM, consistent with predictions based on theoretical thermodynamic data^{24–27}. Initial ethene production was also observed after injection of ethane at 350 °C, consistent with reaction (1) (Table 1). A systematic approach to a constant ethane/ethene concentration ratio was not observed at 350 °C, however, owing to large decreases in ethene concentration with continued reaction (Table 1). This suggests competing reactions such as oxidation and/or reduction of ethene to form CO₂ and CH₄ were occurring at sufficiently rapid rates to preclude equilib-

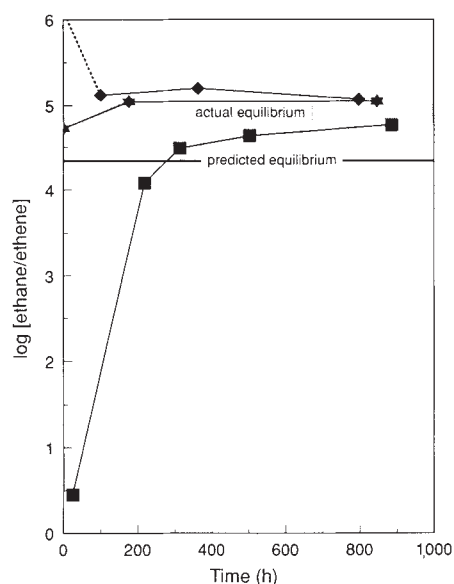


FIG. 2 Variations in the log(ethane/ethene) concentration ratio during reaction with a pyrite–pyrrhotite–magnetite mineral redox buffer at 325 °C and 350 bar after injecting ethene and H₂O (squares), injecting ethane and H₂O (diamonds), and reducing temperature from 350 to 325 °C (stars). The actual equilibrium value bracketed by the experimental data is indicated by the shaded area. The predicted equilibrium value is calculated from compilations of thermodynamic data for aqueous species, minerals and gases^{24–27}.

rium with ethane. Further evidence for this interpretation is provided by decreased rates of CO₂ and CH₄ production in the later stages of experiment 1 that were characterized by relatively low ethene concentrations (Table 1). At 325 °C, however, competing reactions resulting in the degradation of ethene are slow relative to reaction (1) permitting ethene-ethane equilibrium to be established. Indeed, on a return to 325 °C (experiment 5) the aqueous ethene concentration decreased rapidly and the log[ethane/ethene] value increased to 5.05, almost identical to the equilibrium values observed during the other 325 °C experiments (Table 1, Figs 1c and 2).

Although ethane and ethene attained an equilibrium state according to reaction (1) at 325 °C, thermodynamic calculations indicate that ethane and ethene were not in thermodynamic equilibrium with the entire chemical system. Observed ratios of both ethane and ethene to CO₂ and CH₄ deviate by several orders of magnitude from those predicted to be in equilibrium with PPM at 325 °C and 350 bar (refs 24–27). Similarly, equilibrium between dissolved CO₂, CH₄, and PPM was not established during the experiments^{24–27}. The failure of some species to reach equilibrium at 325 °C indicates that kinetic barriers play an important role in controlling fluid chemistry on the timescale of these experiments.

The role of water during oil and gas generation has received much attention in recent years^{17,19,22,28–31}. The experiments presented here demonstrate that the ethane/ethene concentration ratio at 325 °C and 350 bar is regulated by the activity of dissolved H₂, which responds systematically to the redox state of the fluid as a result of the rapid disproportionation of water (reaction (3)). Consequently, ethane and ethene are also in thermodynamic equilibrium with water which represents the source of requisite hydrogen in reaction (1). Water is an effective and continuous source of H₂ during these experiments because the PPM redox buffer consumes oxygen generated by reaction (3). Pyrite, pyrrhotite and magnetite are common minerals in sedimentary basins, but may not coexist to form a redox buffer. Buffered oxygen fugacity, however, is not a prerequisite for water as a source of H₂ because pyrite, pyrrhotite and magnetite all contain reduced iron and will readily consume oxygen individually during oxidation. Moreover, aluminosilicate minerals such as chlorite and smectite contain reduced iron, and are abundant in many sedimentary environments. Other redox-sensitive species may also consume oxygen produced by the disproportionation of water. Lewan²² has suggested that during hydrous pyrolysis of Woodford Shale, water-derived oxygen is consumed via the oxidation of carbonyl groups in kerogen or bitumen to form CO₂.

That water may contribute hydrogen directly to hydrocarbon formation has important implications for oil and natural-gas generation because water is abundant in most sedimentary environments. Thus the initial atomic H/C composition of sedimentary organic matter may not limit the availability of requisite hydrogen for petroleum and natural-gas formation, especially at high temperatures in the deepest parts of sedimentary basins. Water may be particularly important during natural-gas generation owing to the substantial increase in the H/C ratio of CH₄ relative to the source organic material. Consequently, models that do not include water as a source of hydrogen during hydrocarbon formation may significantly underestimate the gas generation potential of organic matter.

The results presented here indicate that the relative abundance of ethane and ethene in laboratory experiments is regulated by reversible equilibrium involving water and redox-sensitive inorganic minerals. This information should be included in models used to constrain the stability of organic species in numerous geological environments, and enhance our understanding of natural alteration processes involving both organic and inorganic sediment components. □

Received 3 January; accepted 13 June 1994.

- Hennet, R. J. C., Crerar, D. A. & Schwartz, J. *Econ. Geol.* **83**, 742–764 (1988).
- Seewald, J. S., Seyfried, W. E. Jr & Shanks, W. C. *Geochim. cosmochim. Acta* (in the press).
- Surdam, R. C. & Crossey, L. J. *Phil. Trans. R. Soc. A* **315**, 135–156 (1985).
- Hunt, J. M. *Petroleum Geochemistry and Geology* (Freeman, San Francisco, 1979).
- Tissot, B. P. & Welte, D. H. *Petroleum Formation and Occurrence* (Springer, New York, 1984).
- Helgeson, H. C., Knox, A. M., Owens, C. E. & Shock, E. L. *Geochim. cosmochim. Acta* **57**, 3295–3339 (1993).
- Shock, E. L. *Geology* **16**, 886–890 (1988).
- Shock, E. L. & Shulte, M. D. *Nature* **343**, 728–731 (1990).
- Shock, E. L. *Origins Life Evol. Biosph.* **20**, 331–367 (1990).
- Kawka, O. E. & Simoneit, B. R. T. *Org. Geochem.* **11**, 311–328 (1987).
- Kvernøden, K. A., Rapp, J. B. & Hostettler, F. D. *Appl. Geochem.* **5**, 83–91 (1990).
- Simoneit, B. R. T. *Can. Mineral.* **26**, 827–840 (1988).
- Simoneit, B. R. T. *Proc. ODP Sci. Res.* (in the press).
- Whiticar, M. J., Faber, E., Whelan, J. K. & Simoneit, B. R. T. *Proc. ODP Sci. Res.* (in the press).
- Quigley, T. M. & Mackenzie, A. S. *Nature* **333**, 549–552 (1988).
- Price, L. C. *Geochim. cosmochim. Acta* **57**, 3261–3280 (1993).
- Monthieux, M., Landais, P. & Monin, J. C. *Org. Geochem.* **8**, 275–292 (1985).
- Seewald, J. S. & Whelan, J. K. *EOS* **74**, 669 (1993).
- Tannenbaum, E. & Kaplan, I. R. *Nature* **317**, 708–710 (1985).
- Tannenbaum, E. & Kaplan, I. R. *Geochim. cosmochim. Acta* **49**, 2589–2604 (1985).
- Mango, F. D., Hightower, J. W. & James, A. T. *Nature* **368**, 536–538 (1994).
- Lewan, M. D. *Am. chem. Soc., Div. Fuel Chem.* (preprints) **37**(4) 1643–1649 (1992).
- Seyfried, W. E. Jr, Janecky, D. R. & Berndt, M. E. in *Experimental Hydrothermal Techniques* (eds Ulmer, G. C. & Barnes, H. L.) 216–240 (Wiley, New York, 1987).
- Kishima, N. *Geochim. cosmochim. Acta* **53**, 2143–2155 (1989).
- Kishima, N. & Sakai, H. *Geochim. J.* **18**, 19–29 (1984).
- Toulmin, P. & Barton, P. B. *Geochim. cosmochim. Acta* **28**, 641–671 (1964).
- Johnson, J. W., Oelkers, E. H. & Helgeson, H. C. *Computers Geosci.* **18**, 899–947 (1992).
- Hoering, T. C. *Org. Geochem.* **5**, 267–278 (1984).
- Lewan, M. D. in *Organic Geochemistry* (eds Engel, M. H. & Macko, S. A.) 419–442 (Plenum, New York, 1993).
- Lewan, M. D. *Proc. 13th World Petroleum Congr.* 1991 Vol. 2 215–223 (Wiley, Chichester, 1993).
- Lewan, M. D., Winters, J. C. & McDonald, J. H. *Science* **203**, 896–899 (1979).

ACKNOWLEDGEMENTS. P. Saccoccia, J. Whelan, T. Eglington and J. Hunt provided comments that improved this manuscript. This research was supported by the US DOE and the US NSF.

In situ evidence for the nature of the seismic layer 2/3 boundary in oceanic crust

R. Detrick, J. Collins, R. Stephen & S. Swift

Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

THE igneous oceanic crust is typically thought of as comprising two layers: an upper crust ('seismic layer 2') characterized by a rapid increase in seismic velocity with depth, and a thicker lower crust ('seismic layer 3') which is distinguished from layer 2 by both a higher P-wave velocity ($6.69 \pm 0.26 \text{ km s}^{-1}$) and a much smaller vertical velocity gradient ($<1 \text{ km s}^{-1} \text{ km}^{-1}$)¹⁻³. A direct correlation has never been established between this seismic layering and the *in situ* lithological and physical properties of oceanic crust. The transition between seismic layers 2 and 3 has been variously interpreted as a change in igneous rock texture from doleritic sheeted dykes to gabbro^{4,5}, an increase in metamorphic grade from greenschist- to amphibolite-facies rocks^{2,6-9}, or a change in bulk crustal porosity with depth^{2,10}. We have re-examined available seismic refraction data from around Hole 504B, the deepest ($>1.8 \text{ km}$) continuous hole drilled into the oceanic crust¹¹⁻¹³, and find that at this location the seismic layer 2/3 boundary lies within the sheeted-dyke complex, where it is associated with gradual downhole changes in crustal porosity and alteration, not a lithological transition from sheeted dykes to gabbro.

Hole 504B, located on the southern flank of the Costa Rica rift in the eastern equatorial Pacific, provides the best opportunity to establish directly the geological nature of the boundary between seismic layers 2 and 3 in the oceanic crust. Drilling on seven different Deep Sea Drilling Project (DSDP) and Ocean Drilling Program (ODP) legs since 1979 has progressively deepened this hole to 2,111 m below the sea floor and $>1,800 \text{ m}$ into the igneous crust, placing it near the expected depth of the layer 2/3 boundary¹¹⁻¹³. The (igneous) crustal section at Hole 504B consists of 571 m of pillow lavas overlying a 200-m-thick transi-

tion zone of mixed pillows and dykes, which is in turn underlain by a sheeted-dyke section extending to the base of the hole (Fig. 1). Changes in the composition and alteration of the dykes, first noted in the section drilled on ODP Leg 111, suggest that the hole may have penetrated the hydrothermal reaction zone presumed to exist near the base of the sheeted-dyke section, but no gabbros have yet been recovered^{12,13}.

A comprehensive suite of logs and borehole experiments have been carried out in Hole 504B (ref. 14), most recently on ODP Leg 148 (ref. 15) (Fig. 1). Sonic velocity logs show that compressional-wave (P-wave) velocities in the upper 600 m of the crustal section are relatively low ($\sim 5 \text{ km s}^{-1}$) and quite variable, ranging from 4 to 6 km s^{-1} . Log-derived crustal resistivities are also low ($\sim 10 \Omega \text{ m}$) indicating high bulk crustal porosities. In the transition zone and upper part of the sheeted-dyke section, sonic-log velocities increase to $>6 \text{ km s}^{-1}$ at $\sim 1,000 \text{ m}$ depth. (Unless otherwise stated, quoted depths are 'sub-basement'—that is, measured from the top of the igneous crustal section.) In this interval bulk crustal porosities decrease rapidly, as indicated by the much higher resistivity measurements. Below this depth, sonic velocities and crustal resistivity continue to increase but much more slowly. By 1.3–1.4 km depth, sonic velocities reach 6.5 km s^{-1} . Based on these observations, the ODP Leg 148 scientific party tentatively placed the top of seismic layer 3 just above the base of the present hole ($\sim 1.8 \text{ km}$ depth), in the lowermost part of the drilled sheeted-dyke section¹³.

Seismic experiments carried out in and around Hole 504B over the past fifteen years can be used to determine the depth of the change in velocity gradient that defines the transition from seismic layer 2 to layer 3, for comparison with the drilling results. These studies include sonobuoy refraction profiles¹⁵⁻¹⁷, borehole seismic experiments on DSDP Legs 70 (ref. 18) and 92 (ref. 19), a vertical seismic profile on Leg 111 (ref. 11) and a multichannel seismic survey and sonobuoy study of the area around Hole 504B (ref. 20). Individually, none of these studies provides a good constraint on the depth to the layer 2/3 boundary. From these published velocity–depth functions, the top of seismic layer 3 may be as shallow as 1.2 km (ref. 16) or as deep as $>2 \text{ km}$ (refs 15, 18). This has led to considerable uncertainty as to whether Hole 504B has reached seismic layer 3. By combining the best constraints from each of these experiments, however we

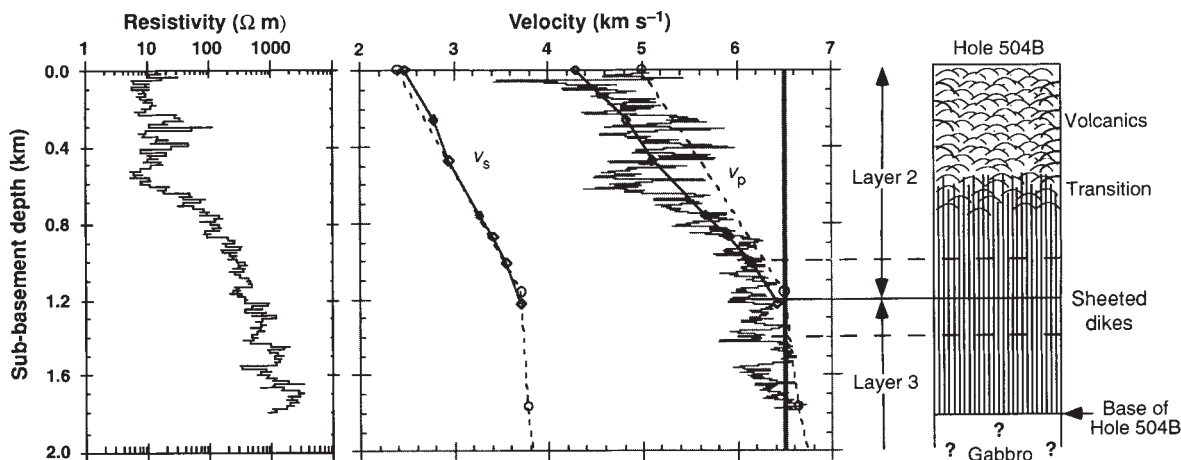


FIG. 1 Summary of (from right to left) the lithostratigraphy in Hole 504B¹¹⁻¹³, the crustal velocity structure at the drill site determined in this study, and the variation in bulk resistivity downhole derived from logging on Leg 148¹³. The crustal velocity model is based on a layer 3 velocity of 6.5 km s^{-1} derived from the sonobuoy studies around Hole 504B (ref. 16) (vertical grey bar) and a linear extrapolation of the shallow crustal velocity gradient determined from travel-time and amplitude modelling of a borehole seismic experiment¹⁹ (thick solid

line). The thin grey line shows the sonic velocity log obtained on Leg 148 for comparison¹³. Also shown is the sonobuoy velocity model of Collins *et al.*¹⁶ (dashed line). We place the change in vertical velocity gradient that defines the seismic layer 2/3 boundary at $1.2 \pm 0.2 \text{ km}$ sub-basement, within the sheeted-dyke section $\sim 600 \text{ m}$ above the base of the present hole. At this location, at least the upper 400–800 m of seismic layer 3 consists of dolerites and metadolerites rather than gabbro.

can obtain a much more accurate estimate of the depth to the layer 2/3 boundary than is possible using any single dataset.

The best constraints on the velocity structure of the shallow crust at Hole 504B are based on travel-time and amplitude analysis of borehole seismic experiments^{18,19}. The results of these experiments substantially agree with one another and show P-wave velocities increasing from $\sim 4.3 \text{ km s}^{-1}$ in the uppermost crust to $>6 \text{ km s}^{-1}$ at $\sim 1 \text{ km}$ depth. Little and Stephen¹⁹ report velocities of 6.4 km s^{-1} at 1.2 km depth; Stephen and Harding¹⁸ extrapolate to velocities of $\sim 6.8 \text{ km s}^{-1}$ at $\sim 2.3 \text{ km}$, with a significant change in velocity gradient at $\sim 1 \pm 0.2 \text{ km}$. In both experiments, the shooting lines did not extend to large enough ranges to identify deeper changes in velocity gradient that might be associated with the layer 2/3 boundary. But the crustal velocity structure down to $1\text{--}1.2 \text{ km}$ depth is very well constrained by these two experiments, which show that P-wave velocities of $>6.4 \text{ km s}^{-1}$ are reached at depths as shallow as 1.2 km .

The main constraint on the deeper crustal structure at Hole 504B comes from sonobuoys deployed within a few tens of kilometres of the drill site^{15–17}. The limitations of sonobuoys for crustal structure studies are well-known; the receivers drift during the course of an experiment, which in the presence of significant basement topography and lateral heterogeneity in crustal structure can lead to large uncertainties in crustal velocity solutions. Another important limitation of some previous sonobuoy seismic studies at Hole 504B is that only first-arrival travel times were modelled^{15,17}. As a result, published sonobuoy velocity solutions around Hole 504B display a range of layer 2 velocities, and estimated depths to the layer 2/3 boundary, which range from 1.2 km to $>4 \text{ km}$ (refs 15–17). The best-constrained solution¹⁶, based on both travel-time and amplitude modelling, places the top of seismic layer 3 at $\sim 1.2 \text{ km}$ depth, well above the base of the present hole. But, the strongest constraint provided by these sonobuoy studies is on the P-wave velocity of layer 3 around Hole 504B. The phase velocity of layer 3 P-wave arrivals is quite uniform, with an average value of $\sim 6.5 \text{ km s}^{-1}$ (ref. 16). This is less than the velocity of 6.7 km s^{-1} typically associated with seismic layer 3, but is still within the normal range of layer 3 velocities^{1,3,5,21}.

Borehole seismic experiments in Hole 504B thus provide accurate estimates of the shallow ($<1 \text{ km}$) crustal velocity structure, but do not constrain layer 3 velocities nor the depth to the layer 2/3 boundary. Sonobuoy seismic studies, on the other hand, do not resolve the shallow crustal velocity structure well, but provide an excellent determination of the velocity of seismic layer 3. We can combine the best-constrained aspects of these datasets to derive a new estimate of the upper-crustal velocity structure at Hole 504B (Fig. 1). We place the layer 2/3 boundary at $1.2 \pm 0.2 \text{ km}$ —equivalent to $\sim 1,500 \text{ m}$ below the sea floor. This estimate is based on a linear extrapolation of the shallow crustal velocity gradient determined by Little and Stephen¹⁹ to the depth where velocities reach the 6.5 km s^{-1} value known to be typical of seismic layer 3 around Hole 504B¹⁶. The error of $\pm 200 \text{ m}$ assigned to this depth is a qualitative estimate based on the uncertainty in the velocity gradient determined by Little and Stephen¹⁹ using both travel-time and amplitude modelling, and the agreement between this velocity structure and the downhole changes in sonic velocity shown in Fig. 1.

Our results thus suggest that the seismic layer 2/3 boundary is located more than 600 m above the base of Hole 504B near the middle of the drilled sheeted-dyke section. At this depth sonic velocities are close to the 6.5 km s^{-1} value typical of layer 3 refractors in this area, and P-wave velocity gradients in the underlying section are less than $0.5 \text{ km s}^{-1} \text{ km}^{-1}$. If our correlation is correct, the seismic layer 2/3 boundary at Hole 504B is not a lithological transition and at least the upper $400\text{--}800 \text{ m}$ of seismic layer 3 consists of fine-grained, massive dolerites and metadolerites rather than gabbro. The depth interval in which we place the layer 2/3 boundary in Hole 504B coincides with a

progressive change in the character of the sheeted dykes that was first observed in the section drilled on ODP Leg 111 at $\sim 1.3 \text{ km}$ sub-basement¹¹, and was also seen in the rocks subsequently drilled on Legs 140 (ref. 12) and 148 (ref. 13). The lowermost 500 m of the sheeted-dyke section is associated with an increase in grain size and a scarcity of chilled, glassy dyke margins, which suggest higher temperatures and slower cooling rates^{12,13}. The alteration character of the dykes also progressively change downhole as chlorite abundances decrease, the proportion of amphibole increases, the composition of the amphibole changes from mainly actinolite in the upper dyke section to more hornblende-rich amphibole, and secondary anorthite begins to appear^{12,13,22}. All of these changes are quite gradual; the seismic layer 2/3 boundary in Hole 504B is not marked by sharp downhole changes in grain size, crustal porosity, degree of alteration or igneous stratigraphy.

The results reported here are consistent with several independent lines of evidence that suggest that the seismic layer 2/3 boundary is not a simple stratigraphic contact between sheeted dykes and gabbro. Evidence has accumulated since the 1970s that the seismic structure of the shallow ($<2 \text{ km}$) oceanic crust primarily reflects variations in bulk porosity rather than igneous rock type^{2,10,23,24}. This has led to the suggestion that the top of layer 3 may be controlled, at least in part, by the depth at which cracks become effectively closed by lithostatic pressure^{2,10}. Variations in metamorphic grade and mineralogy with depth, first noted by Christensen⁶, have been proposed by a number of investigators as a cause of the change in seismic velocity at the layer 2/3 boundary^{2,4,8,9,25}. On the basis of correlations between oceanic refraction results and laboratory measurements of the physical properties of ophiolite samples, the layer 2/3 boundary has been placed within the sheeted-dyke section in at least five ophiolite complexes, including the Troodos²⁶, Bay of Islands^{7,9} and Oman²⁷ ophiolites. These results are consistent with the view that the upper portion of seismic layer 3 consists predominantly of metadolerite, rather than gabbro^{2,7}. Swift and Stephen²⁸ also argued that gabbro is not an important constituent of at least the uppermost part of seismic layer 3, from the low Q values (high attenuation) measured for gabbros from ODP Hole 735B and ophiolites.

Taken together, these lines of evidence, including our new results from Hole 504B, indicate that the seismic layer 2/3 boundary in oceanic crust should be viewed as a transition in physical properties occurring over a depth interval of several hundred metres. This transition controlled by many factors including bulk porosity, fracture geometry and crustal alteration, as well as by igneous lithology^{2,5,23}. The progressive downhole changes in crustal porosity and alteration that define the layer 2/3 transition in Hole 504B may either be a primary feature of the crust (thus leading to lower water/rock ratios and higher-temperature alteration of this part of the section) or a result of the alteration process itself. In the latter case, the change in porosity gradient that defines the layer 2/3 boundary may not be fixed in depth but may migrate up-section with the progressive alteration of the lower sheeted-dyke section and infilling of cracks by alteration products as the crust ages. The stratigraphic position of the seismic layer 2/3 boundary relative to the gabbro section may also depend on the depth of the axial magma chamber and thus the thickness of the sheeted-dyke section. Where magma chambers are shallow ($\sim 1\text{--}1.5 \text{ km}$ depth), as along most fast-spreading ridges^{29,30}, porosities may be relatively high throughout most of the sheeted-dyke section, and the top of seismic layer 3 may be located near the dyke/gabbro transition. But where crustal magma chambers are deeper ($>2 \text{ km}$), as along some intermediate and slow-spreading ridges^{31,32}, bulk porosities in the lower part of the thick sheeted-dyke section may be lower, higher-temperature alteration more pervasive, and the top of seismic layer 3 may occur far above the dyke/gabbro transition. Thus, although the boundary between seismic layers 2 and 3 may coincide with an igneous contact in some locations

(or at some crustal ages), the available data indicate that this seismic boundary should not generally be interpreted as a simple lithological transition from dykes to gabbro. □

Received 4 March; accepted 20 June 1994.

1. Raitt, R. W. in *The Sea* Vol. 3 (ed. Hill, M. N.) 85–102 (Wiley-Interscience, New York, 1963).
2. Spudich, P. & Orcutt, J. A. *Rev. Geophys. Space Phys.* **18**, 627–645 (1980).
3. White, R. S., McKenzie, D. & O'Nions, R. K. *J. geophys. Res.* **97**, 19683–19716 (1992).
4. Fox, P. J., Schreiber, E. & Peterson, J. J. *J. geophys. Res.* **78**, 5155–5172 (1973).
5. Fox, P. J. & Stroup, J. B. in *The Sea* (ed. Emiliani, C.) 119–218 (Wiley & Sons, New York, 1981).
6. Christensen, N. I. *Geol. Soc. Am. Bull.* **81**, 905–908 (1970).
7. Peterson, J. J., Fox, P. J. & Schreiber, E. *Nature* **247**, 194–196 (1974).
8. Christensen, N. I. & Salisbury, M. H. *Rev. Geophys. Space Phys.* **13**, 57–86 (1975).
9. Salisbury, M. H. & Christensen, N. I. *J. geophys. Res.* **83**, 805–817 (1978).
10. Lort, J. M. & Matthews, D. H. *J. R. astr. Soc.* **27**, 383–392 (1972).
11. Becker, K. et al. *Init. Rep. ODP Leg 111 (Part A)* 35–251 (1988).
12. Dick, H. J. B. et al. *Init. Rep. ODP Leg 140* 37–200 (1992).
13. Alt, J. C. et al. *Init. Rep. ODP Leg 148* 27–121 (1993).
14. Becker, K. et al. *Rev. Geophys.* **27**, 79–102 (1989).
15. Hobart, M. A., Langseth, M. G. & Anderson, R. N. *Init. Rep. DSDP Leg 83* 379–399 (1985).
16. Collins, J. A., Purdy, M. G. & Brocher, T. M. *J. geophys. Res.* **94**, 9283–9302 (1989).
17. Zehnder-Mutter, C. *JOI/USSAC Newsletter* **4** (1992).
18. Stephen, R. A. & Harding, A. J. *J. geophys. Res.* **88**, 8289–8298 (1983).
19. Little, S. A. & Stephen, R. A. *Init. Rep. DSDP Leg 83* 517–528 (1985).
20. Collins, J. A., Brocher, T. M. & Purdy, G. M. *Proc. ODP Sci. Res. Leg 111* 177–191 (1989).
21. Shor, G. G. Jr, Menard, H. W. & Raitt, R. W. in *The Sea* Vol. 4 (ed. Maxwell, A. E.) 3–27 (Wiley-Interscience, New York, 1970).
22. Laverne, C. et al. *Proc. ODP, Sci. Res. Leg 140* (in the press).
23. Casey, J. F., Dewey, J. F., Fox, P. J., Karson, J. A. & Rosencrantz, E. in *The Sea* Vol. 7 (ed. Emiliani, C.) 305–338 (Wiley-Interscience, New York, 1981).
24. Becker, K. et al. *Nature* **300**, 594–598 (1982).
25. Carlson, R. L. & Herrick, C. N. *J. geophys. Res.* **95**, 9153–9170 (1990).
26. Christensen, N. I. *Tectonophysics* **47**, 131–157 (1978).
27. Christensen, N. I. & Smewing, J. D. *J. geophys. Res.* **86**, 2545–2556 (1981).
28. Swift, S. & Stephen, R. *Geophys. Res. Lett.* **19**, 1871–1874 (1992).
29. Detrick, R. S. et al. *Nature* **326**, 35–41 (1987).
30. Detrick, R. S. et al. *Science* **259**, 499–503 (1993).
31. Morton, J. L., Sleep, N. H., Normark, W. R. & Tompkins, D. H. *J. geophys. Res.* **92**, 11315–11326 (1987).
32. Collier, J. & Sinha, M. *Nature* **346**, 646–648 (1990).

ACKNOWLEDGEMENTS. We thank H. J. B. Dick for stimulating our interest in the depth to layer 3 at Hole 504B. We also thank K. Gillis, J. Alt, P. J. Fox, C. Laverne, M. Salisbury, R. Carlson and J. Orcutt for their critical reviews and insightful comments, and the scientists and engineers who have participated in the drilling at Hole 504B. This work was supported by the US NSF and ODP.

Species coexistence and self-organizing spatial dynamics

Michael P. Hassell*, Hugh N. Comins† & Robert M. May‡

* Department of Biology and NERC Centre for Population Biology, Imperial College, Silwood Park, Ascot, Berkshire SL5 7PY, UK

† Department of Biological Sciences, University of New South Wales, PO Box 1, Kensington, NSW 2033, Australia

‡ Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

IN a patchy environment, dispersal between neighbouring local populations can allow the total (regional) population to persist^{1–5}, even where all patches are identical and the within-patch dynamics are unstable, the total population readily persists as a metapopulation. This persistence is associated with striking, self-organized spatial patterns in the densities of the subpopulations. In the case of hosts and parasitoids, these may form spiral waves, spatial chaos, or a so-called 'crystal lattice' with regularly spaced knots of high population density^{4,6}. Here we extend earlier work on two species to three or more, showing that coexistence of competing species is usually associated with some degree of persistent spatial segregation, even when the environment is uniform. At its most extreme, this can confine one species to small, relatively static 'islands' within the habitat, giving the appearance of isolated pockets of favourable habitat. The distributions of interacting species may thus result from a trade-off between dispersal and competition within subpopulations, as much as from external factors.

Insect host–parasitoid interactions make convenient systems for study. In a homogeneous environment, the dynamics of a two-parasitoid–one-host interaction in which there are discrete generations and both parasitoids search simultaneously for the same host stage may be described by

$$N_{t+1} = \lambda N_t f(P_t + rQ_t)$$

$$P_{t+1} = cN_t[1 - f(P_t + rQ_t)][P_t/(P_t + rQ_t)]$$

$$Q_{t+1} = dN_t[1 - f(P_t + rQ_t)][rQ_t/(P_t + rQ_t)]$$

Here N , P and Q are, respectively, the host and two parasitoid population sizes in generations t and $t + 1$, λ (assumed for simplicity to be density independent) is the host rate of increase, the function f is the probability of a host not being found by either the P_t or Q_t searching parasitoids, r is the searching efficiency of the parasitoid Q relative to parasitoid P , and c and d , respectively, define the average number of adult female parasitoids of P and Q emerging from each host attacked.

We now assume that the habitat is divided into a large number of discrete patches (for example, food plants for a herbivorous insect) in which the adult hosts oviposit, their larvae feed and adult parasitoids search for immature hosts to parasitize. Further, we assume that the patches are identical and arranged in a two-dimensional grid of identical patches, assumed in the present study to be 50×50 patches in size. Only at much smaller scales are our findings affected by the choice of grid size (see below). Each model generation has two phases: (1) interactions of hosts and parasitoids within individual patches and (2) dispersal between patches. In the first phase, the interaction of the populations within each patch is described by the basic Nicholson and Bailey equations^{7,8} for parasitoids which search independently and randomly, without any of the usual additional features that, for example, complicate the parasitoid functional response^{9,10}, make the host growth rate density-dependent^{11,12}, or implicitly represent a heterogeneous environment in a phenomenological way¹³. Thus, within a patch i , $f = \exp(-a_1P_i - a_2Q_i)$, where a_1 and a_2 are the per capita searching efficiencies of the P_i and Q_i searching parasitoids, respectively. This is done to ensure that the within-patch dynamics will always be locally unstable and that any metapopulation persistence results expli-

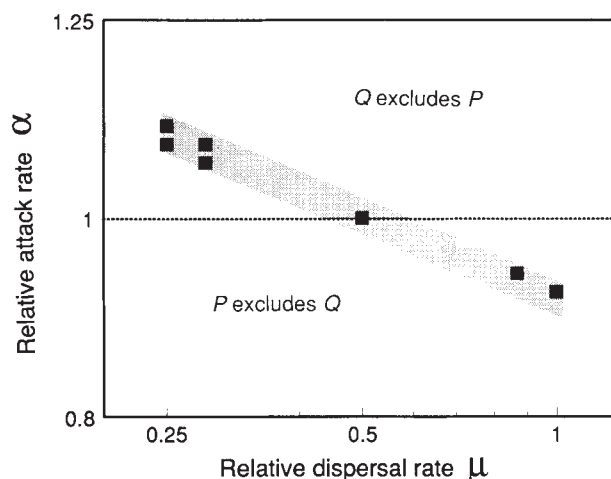


FIG. 1 Simulation outcomes of the one-host–two-parasitoid model with $\lambda = 2$, $\mu_P = 0.5$, $\mu_N = 0.5$, for a logarithmically spaced grid of relative dispersal rate μ and relative attack rate α (a measure of relative competitive ability), distributed around the two-identical-parasitoids case ($\mu = \alpha = 1$). The regions in which parasitoid species Q excludes species P , and in which P excludes Q are indicated. These are separated by an imprecise region (shaded) in which coexistence was observed for the parameter combinations marked by black dots. The horizontal broken line separates the regions where Q is competitively superior (above) from P being competitively superior (below).

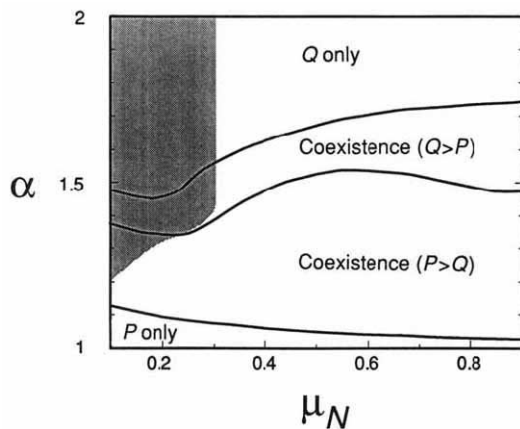


FIG. 2 Simulation outcomes of the one-host-two-parasitoid model with $\lambda=2$, for a linear grid of host dispersal rates μ_N and relative attack rate α . Parasitoid species Q is the less dispersive parasitoid ($\mu_P=0.2$, $\mu_Q=0.02$). The regions separated by solid lines are, from top to bottom: species Q (low disperser) excludes species P , coexistence with species Q more numerous than P , coexistence with species P more numerous than Q , species P (high disperser) excludes species Q . The shaded area denotes the region where chaotic spatial patterns occur, while the remaining parameter combinations give spirals.

citly from the spatially distributed nature of the interaction. The single-patch case is thus always unstable. In the second, dispersal phase, we allow for dispersal by a certain fraction of the new adult hosts (μ_N) and adult parasitoids (μ_P and μ_Q) in each patch redistributing themselves to the eight immediately neighbouring patches, while the remainder stay behind to reproduce in their native patch. The dispersal is spatially symmetrical for all three species, but the fraction of dispersers is species-dependent.

Coexistence of all three species in this system occurs readily within a subset of the parameter combinations that allow persistence in the equivalent two-species spatially distributed models⁶. Also, as in the two-species case, the spatial patterns of local population abundance underlying this persistence take the form of spiral waves or chaotic patterns. We quantify the relative competitive advantages of the two parasitoid species using two ratios: α (defined as a_2d/a_1c) for the relative parasitoid efficiency at finding hosts and converting them into adult female parasitoids in the next generation, and μ (defined as μ_Q/μ_P) for the relative dispersal rates of the two parasitoid species. Thus, for two identical parasitoids, $\mu=\alpha=1$. Figure 1 summarizes the results from a number of simulations in which μ is varied around the region of $\alpha=1$ (that is, the parasitoids have very similar attack rates). Competitive exclusion is the norm, with the winning species being determined by the balance of the advantages conferred by relatively fast dispersal and by relatively high searching efficiency. Coexistence is only observed (dots in Fig. 1) very close to the boundary separating the regions of exclusion of one or the other parasitoid species. Note that, in the 'south-east' sector of Fig. 1, Q and P coexist even though Q is both

competitively and dispersively inferior to P . Such coexistence is not found in simpler metapopulation models which trade competitive against dispersal abilities¹⁴.

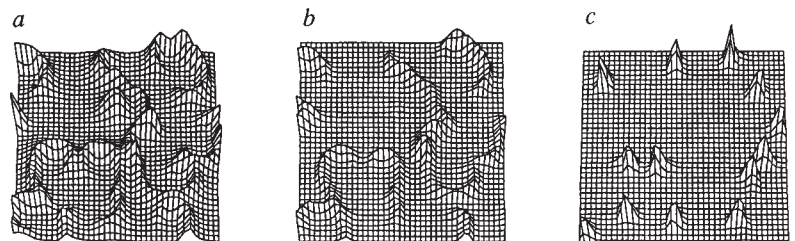
The picture becomes different if the dispersal rates of the two parasitoid species differ more markedly; coexistence can now occur for a wide range of relative attack rates (Fig. 2). Furthermore, we find that this coexistence tends to be associated with a marked degree of self-organizing spatial separation between the competing species. This new and unexpected result is seen at its most extreme in the case of spiral-wave dynamics for two competing parasitoids with very different dispersal rates. In this case, the less dispersive species tends to be confined to the central foci of the spirals where it may be the more abundant species, and the more dispersive species occupies the remainder of the 'trailing arm' of the spirals (Fig. 3). Because the foci of spirals are relatively static in these models, the less dispersive species seems to occur only in isolated, small 'islands' within the habitat, much as if these were pockets of favourable habitat. It can thus persist with total density very much lower than that of the more dispersive species. As the dispersal rates of the two species become more similar, or the relative attack rate of the less dispersive species is increased, it comes to dominate progressively larger areas around the foci; in other words, the apparent spatial 'niche' of the less dispersive species spreads farther into the arm of the spirals.

Coexistence is also observed (albeit for a more restricted range of parameter combinations) when the dynamics is spatial chaos rather than spirals. Persistence in these cases can again be understood in terms of dynamically induced spatial segregation, but with the favourable locations being transient rather than static. In these cases, the less dispersive species cannot coexist at very low relative total density.

The above results also apply to two other three-species models that we have investigated: one with two hosts and one parasitoid and the other a host-parasitoid-hyperparasitoid interaction (H.N.C. and M.P.H., manuscript in preparation). These two models give remarkably similar results to the one-host-two-parasitoid model. In each case it is possible for a third, relatively poorly dispersing species to coexist stably within the self-organized spatial structures generated by the basic two-species host-parasitoid interaction. The third species can be another host or a hyperparasitoid. In the two-host-one-parasitoid system the low-dispersing host must have a higher rate of increase, or be less strongly attacked by the parasitoid. In the host-parasitoid-hyperparasitoid system, the low-dispersing hyperparasitoid must have a higher searching efficiency than the parasitoid (this agrees with the conclusions from non-spatial models of host-parasitoid-hyperparasitoid interactions¹⁵⁻¹⁷). The same rules seem to apply in more complicated systems; for example, in simulations with four or five species where one of the host and one of the parasitoid species are highly dispersive, and the other two or three species are relatively sedentary.

Thus, our multispecies extension of work on two-species systems shows that local dispersal combined with inherently unstable, oscillatory local population dynamics can lead to self-organized spatial structures—spiral waves or chaos—which

FIG. 3 Maps of the spatial density distribution (with linear scales) of hosts (a), highly dispersive parasitoids (b) and relatively sedentary parasitoids (c), in a snapshot from the dynamics of a persistent parasitoid-host-parasitoid system with $\lambda=2$, $\mu_N=0.5$, $\mu_{P1}=0.5$, $\mu_{P2}=0.05$, $\alpha=1.3$. The grids must be mentally superimposed to perceive the relationships between the densities of the various species. Spiral foci exist at the ends of the 'mountain ranges' in a (excluding ends at the edges of the grid). In the time-evolution of the system the mountain ranges are the peaks of population density waves and are thus in continuous motion. The foci, by contrast, remain in almost exactly the same place, for indefinitely long times. The vertical scale in c is extended 10-fold.



enable different combinations of three or more interacting species to coexist. In the simplest case of one host and two parasitoid species, the competitively inferior parasitoid is favoured at the foci of spiral waves of abundance; in essence, it occupies a spatial niche created by the basic two-species dynamics. The possibility that coexistence, and thus species diversity, is promoted by trade-offs between competitive and dispersal abilities has long been recognized^{18–24}. What is new here is that the spatial metapopulation structure, within which these trade-offs occur, is itself self-generated on an intrinsically homogeneous substrate. For plausible mixes of species traits, we would expect often to find competitors separated in space in a self-organized way, as a result of local dynamics and differences in dispersal rates; such patterns are often seen but are traditionally interpreted as caused by species specialization within spatially heterogeneous environments. There is a need for new tests that can distinguish deterministic and self-generated spatial structure from spatial randomness (paralleling work which tries to distinguish low-dimensional deterministic chaos from 'real noise'^{25,26}). Finally, we emphasize that this kind of coexistence requires the characteristic linear scale of the self-organized spatial structures (for example, the wavelength of the spirals) to be significantly less than the typical linear dimension of the arena occupied by the community in question. If habitat destruction were to bring the community below its critical spatial size (which, being a collective property of the system, is not easily inferred from the movements of individuals), we would expect the system to collapse²⁷. □

Received 24 March; accepted 21 June 1994.

1. Sole, R. V., Bascompte, J. & Valls, J. J. *theor. Biol.* **159**, 469–480 (1992).
2. Swinton, J. & Anderson, R. M. in *Ecology of Infectious Diseases in Natural Populations* (eds Grenfell, B. & Dobson, A.) (Cambridge Univ. Press, UK, 1994).
3. Halley, J. M., Comins, H. N., Lawton, J. H. & Hassell, M. P. *Oikos* (in the press).
4. Hassell, M. P., Comins, H. N. & May, R. M. *Nature* **353**, 255–258 (1991).
5. Sole, R. V. & Valls, J. *Phys. Lett. A* **166**, 123–128 (1992).
6. Comins, H. N., Hassell, M. P. & May, R. M. *J. Anim. Ecol.* **61**, 735–748 (1992).
7. Nicholson, A. J. *J. Anim. Ecol.* **2**, 131–178 (1933).
8. Nicholson, A. J. & Bailey, V. A. *Proc. zool. Soc., Lond.* **3**, 551–598 (1935).
9. Holling, C. S. *Can. Ent.* **91**, 293–320 (1959).
10. Holling, C. S. *Can. Ent.* **91**, 385–398 (1959).
11. Beddington, J. R., Free, C. A. & Lawton, J. H. *Nature* **255**, 58–60 (1975).
12. May, R. M., Hassell, M. P., Anderson, R. M. & Tonkyn, D. W. *J. Anim. Ecol.* **50**, 855–865 (1981).
13. May, R. M. *J. Anim. Ecol.* **47**, 833–843 (1978).
14. Tilman, G. D. *Ecology* **75**, 2–16 (1994).
15. Beddington, J. R. & Hammond, P. S. *J. Anim. Ecol.* **46**, 811–821 (1977).
16. Hassell, M. P. in *Population Dynamics* Vol. 1 (eds Anderson, R. M., Turner, B. D. & Taylor, L. R.) 283–306 (Blackwell, Oxford, 1979).
17. May, R. M. & Hassell, M. P. *Am. Nat.* **117**, 234–261 (1981).
18. Hutchinson, G. E. *Ecology* **32**, 571–577 (1951).
19. Skellam, J. G. *Biometrika* **38**, 196–218 (1951).
20. Levins, R. & Culver, D. *Proc. Natn. Acad. Sci. U.S.A.* **68**, 1246–1248 (1971).
21. Horn, H. S. & MacArthur, R. H. *Ecology* **53**, 749–752 (1971).
22. Hanski, I. & Ranta, E. *J. Anim. Ecol.* **52**, 263–279 (1983).
23. Nee, S. & May, R. M. *J. Anim. Ecol.* **61**, 37–40 (1991).
24. Hanski, I. & Zhang, D. Y. *J. theor. Biol.* **163**, 491–504 (1993).
25. Sugihara, G. & May, R. M. *Nature* **344**, 734–741 (1990).
26. Farmer, J. D. & Sidorowich, J. J. in *Evolution, Learning and Cognition* (ed. Lee, Y. C.) 277–304 (World Scientific, New York, 1989).
27. Hassell, M. P., Godfray, H. C. J. & Comins, H. N. in *Biotic Interactions and Global Change* 402–423 (Sinauer, Sunderland, 1992).

ACKNOWLEDGEMENTS. The work was supported by the Natural Environment Research Council (M.P.H.), the NERC Centre for Population Biology at Silwood Park (H.N.C.), and the Royal Society (R.M.M.).

Speed and cerebral correlates of syllable discrimination in infants

Ghislaine Dehaene-Lambertz
& Stanislas Dehaene

Institute of Cognitive and Decision Sciences, University of Oregon, Eugene, OR 97405, USA and Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS, INSERM and EHESS, 54 Boulevard Raspail, 75270 Paris cedex 06, France

THE remarkable linguistic abilities of human neonates are well documented^{1–5}. Young infants can discriminate phonemes even if they are not used in their native language^{2–4}, an ability which regresses during the first year of life^{4,5}. This ability to discriminate is often studied by repeating a stimulus for several minutes until some behavioural response of the infant habituates, and later examining whether the response recovers when the stimulus is changed⁶. This method, however, does not reveal how fast infants can detect phonetic changes, nor what brain mechanisms are involved. We describe here high-density recordings of event-related potentials in three-month-old infants listening to syllables whose first consonants differed in place of articulation. Two processing stages, corresponding to an increasingly refined analysis of the auditory input, were identified and localised to the temporal lobes. A late frontal response to novelty was also observed. The infant brain recognizes a phonetic change in less than 400 ms.

Sixteen healthy two- to three-month-old infants (average 81 days old, range 63–91), born to monolingual American-English families, were tested. Infants were seated in a carrier affixed to the parent, and their heads were covered with a geodesic sensor net, a very light mesh made of 58 electrodes encased in sponges soaked with a saline solution⁷. Infants faced a loudspeaker placed on top of a TV monitor in a sound-attenuated room. To avoid eye and head movement, a silent video showing attention-grabbing coloured objects was played continuously. The video was not synchronised with the auditory stimuli, thus preventing

any visually evoked potentials from appearing after averaging in synchrony with the auditory stream.

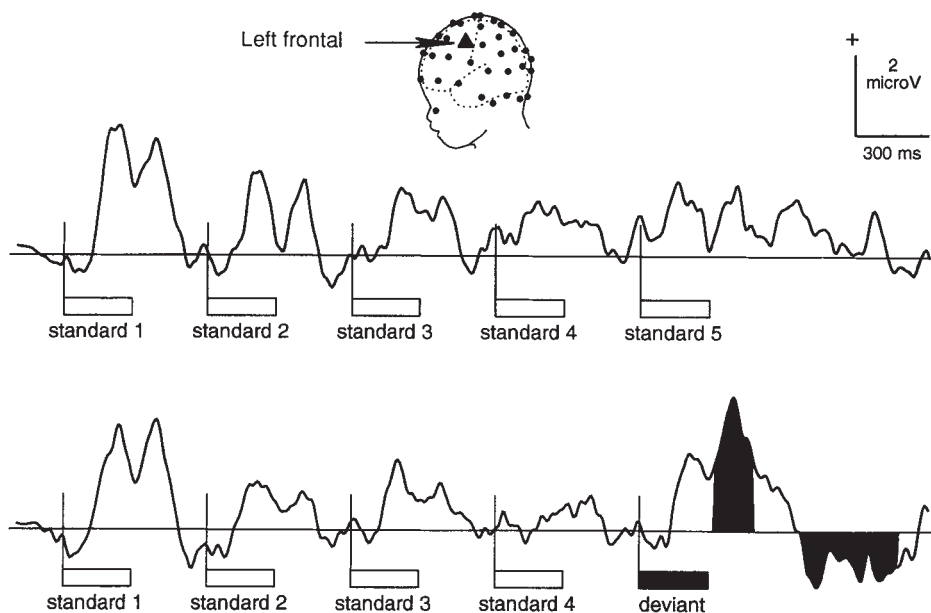
On each trial, a sequence of five syllables (/ba/ or /ga/) was presented. In half the trials, one syllable, designated as the standard, was repeated five times (standard trials). In the other half (deviant trials), the standard was repeated only four times, followed by one instance of the other syllable, designated as the deviant. Because repeated and deviant trials were randomly mixed, infants could not predict the nature of the fifth stimulus. Thus any significant difference in event-related potentials (ERPs) to repeated and deviant trials indicated that the two syllables had been discriminated. The onset and topography of the observed differences were used to infer the speed and brain mechanisms of syllable discrimination. The evolution of ERPs with successive repetitions of the standard was also studied. In human adults, reduced ERPs to repetitive sounds as well as novelty-specific responses have been described^{8–10}.

Each syllable presentation elicited a distinctive waveform characterized by two peaks (Fig. 1). The maturation of these peaks from birth to six months has been previously described¹¹, but their functional significance has not been elucidated. The first identifiable event, peak 1, reached its maximum about 220 ms after stimulus onset. Peak 1 amplitude was highest in response to the first syllable of each trial. As soon as the standard was repeated, a significant decrease was observed. Further repetitions of the standard did not lead to further decrease, and there was no recovery to the deviant syllable (Fig. 2). Thus, by 200 ms, the acoustical analysis had not proceeded far enough to separate two syllables that were quite similar in pitch, duration and intensity. The generators of peak 1 appeared insensitive to the subtle acoustical differences that encoded phonetic information. The scalp topography of peak 1 showed synchronous anterior positivities and posterior negativities with a temporal-central inversion plane (Fig. 3), suggesting bilateral generators inside the temporal lobes. We speculate that peak 1 reflected the activation of primary and secondary auditory areas.

The next identifiable event, peak 2, reached its maximum about 390 ms after stimulus onset. Again, the amplitude of peak 2 decreased significantly between the first and the second presentation of the standard, with no further decrease to subsequent

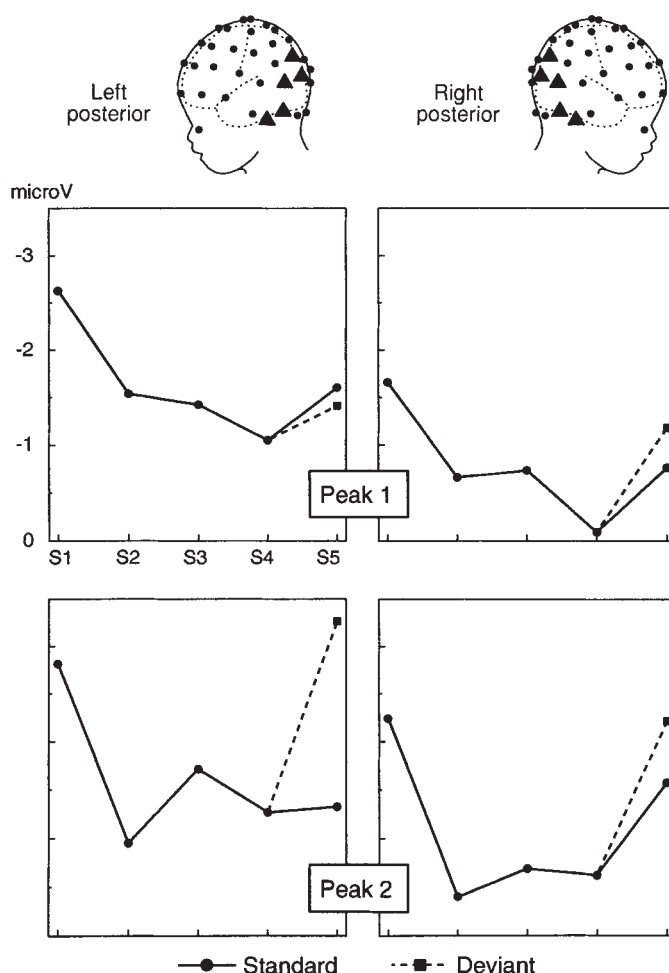
FIG. 1 Grand-averaged ERPs recorded from a left frontal electrode on standard and deviant trials.

METHODS. Two digitized syllables /ba/ and /ga/, naturally produced by a male speaker, were matched for total duration (289 ms), formant transition duration (87 ms), intensity (78 dB) and fundamental frequency (122 Hz). For each infant, one syllable, counterbalanced across subjects, was designated as the standard (S) and the other as the deviant (D). Syllables were presented in groups of five with a stimulus onset asynchrony of 600 ms. On standard trials, the standard was repeated five times (S S S S S) whereas on deviant trials the deviant was introduced in the fifth position (S S S S D). Fifty standard and fifty deviant trials were randomly mixed with an intertrial interval of 3,100 ms. Scalp voltages were recorded from 58 Ag/AgCl electrodes mounted on a Geodesic Sensor Net⁷ applied in anatomical reference to the canthomeathal line and referenced to the right mastoid. Two frontal and two infra-orbital electrodes monitored for eye movements. A ground lead was placed on the seventh cervical vertebra. Voltages were amplified (band pass 0.1–50 Hz, 3 dB per octave attenuation, 60 Hz notch filter), digitized at 125 Hz for 4,096 ms (512 samples) starting 200 ms before trial onset, and automatically edited for eye and motion artefacts. Seven infants with less than 25 artefact-



free trials were rejected. Data from the remaining 16 infants (average 51.2 trials per infant) were averaged, transformed using the average-reference method, baseline corrected and digitally filtered (band pass 0.5–20 Hz).

FIG. 2 Evolution of peak 1 and peak 2 voltages from left and right posterior temporal electrodes across the five standards and the deviant syllable. Voltage was averaged over five adjacent electrodes across two 160 ms time windows centred on each peak. A repeated-measures analysis of variance (ANOVA) was done with peak, stimulus number, hemisphere, and condition (standard versus deviant) as factors, with a Greenhouse–Geisser correction for non-sphericity. The decrease between stimulus 1 and stimulus 2 was significant for both peak 1 ($P=0.0012$) and peak 2 ($P=0.0001$). There was a recovery to the deviant syllable only for peak 2, as measured by the condition $\times$ stimulus number interaction (peak 1, $P=0.36$; peak 2, $P=0.021$) and by the main effect of condition on stimulus 5 (peak 1, $P=0.75$; peak 2, $P=0.0008$). The latter effect was larger over the left hemisphere than over the right (interaction $P=0.014$), and both peaks were of higher amplitude over the left hemisphere than over the right (both P s < 0.011). Another ANOVA on five adjacent superior prefrontal electrodes showed similar decrement and discrimination effects, and a non-significant trend towards left lateralization. In addition, a later time window (680–1,080 ms) showed a significant main effect of condition ($P=0.002$), only deviant stimuli eliciting a bilateral frontal negativity (see Fig. 1).



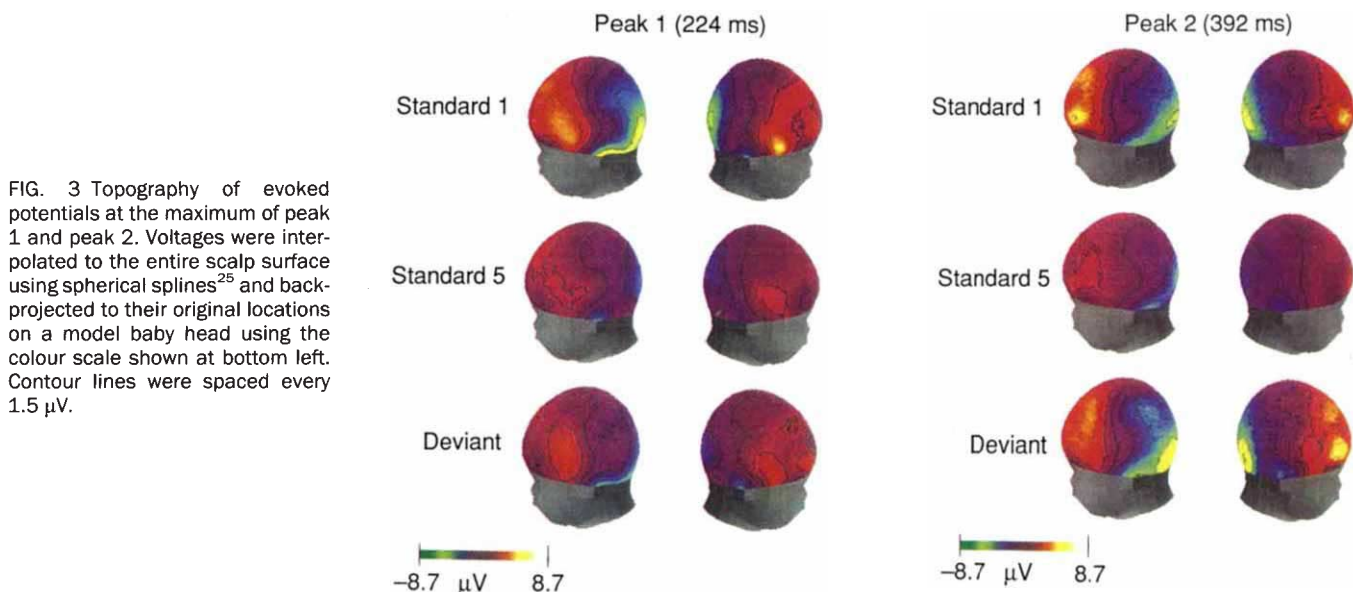


FIG. 3 Topography of evoked potentials at the maximum of peak 1 and peak 2. Voltages were interpolated to the entire scalp surface using spherical splines²⁵ and back-projected to their original locations on a model baby head using the colour scale shown at bottom left. Contour lines were spaced every 1.5 μ V.

repetitions. However, when the deviant syllable was introduced, peak 2 recovered to at least its original level (Fig. 2). Thus, contrary to peak 1, the decrease of peak 2 with repeated stimuli was syllable-specific, and a change in the initial consonant was sufficient to elicit a complete recovery. The neural generators of peak 2 were clearly sensitive to phonetic information, indicating that a single instance of a novel syllable could be recognized in less than 400 ms by the infant brain. Further studies should verify if peak 2 recovery is driven by a categorical perception of phonemes¹, or if it would be elicited by any acoustical change, even within a given phonemic category.

The topography of peak 2, like that of peak 1, showed a polarity inversion along the anterior-posterior axis, suggesting temporal lobe generators. The anterior positivity, however, was more medial for peak 2 than for peak 1 (Fig. 3), confirming that these functionally distinct components originated from distinct brain areas. A forward dipole modelling algorithm¹² suggested that the generators of peak 2 were more posterior, superior and tangential than those of peak 1. Precise localization must remain tentative because of the simplistic assumptions of dipole reconstruction methods.

The anatomy of the infant brain shows hemispheric asymmetries comparable to those found in adults, notably in the temporal lobes¹³⁻¹⁵. Here there were noticeable ERP asymmetries. The amplitudes of peak 1 and peak 2 were significantly larger over left than over right posterior electrodes, as was the size of the recovery of peak 2 to deviant syllables (Fig. 2). This may reflect a functional asymmetry indicating superior processing of short syllables in the left hemisphere. However, ERP summation might also be affected by asymmetries in brain morphology, for instance in the orientation of the left and right sylvian fissures¹⁵. It should be noted that in individual subjects, considerable variability was found in the size and sometimes the direction of asymmetries. This suggests at most a moderate left-hemispheric advantage for language in infants rather than a sharp division of functions. It may explain why previous studies of hemispheric lateralization in infants have not always reported a noticeable left-hemisphere advantage¹⁶⁻¹⁹.

After the fifth syllable of each trial, ERPs were recorded for a longer period of 1,496 ms during which no further stimuli were presented. During this period, a late frontal negativity was observed from 680 ms to 1,080 ms following deviant but not standard syllables (Fig. 1). It showed a widespread topography over the left and right frontal electrodes, with no significant left-right asymmetry. A similar anterior negativity has been observed by others in response to unexpected visual and auditory

stimuli²⁰⁻²³. This suggests that the relevant parameter for eliciting this component is the novelty of the stimuli regardless of their input modality. Thus, two- to three-month-old infants may already possess a supramodal anterior network for novelty detection, perhaps involving arousal and attention-orienting processes^{23,24}, and which can be activated in less than one second.

The fact that young infants can discriminate phonemes and react to novelty was long known from behavioural methods. High-density ERPs, however, go beyond the simple listing of infants' abilities by permitting the decomposition of complex capacities into a series of processing steps, whose duration and brain implementation can be estimated. The method, although still lacking in spatial resolution, should provide new windows into the brain mechanisms and the fine temporal sequence of cognitive development processes. □

Received 23 March; accepted 1 June 1994.

1. Eimas, P. D., Siqueland, E. R., Jusczyk, P. W. & Vigorito, J. *Science* **171**, 303-306 (1971).
2. Best, C. T., McRoberts, G. W. & Nomathemba, M. S. *J. exp. Psych. hum. Percept. Perform.* **13**, 345-360 (1988).
3. Eimas, P. D., Miller, J. L. & Jusczyk, P. W. in *Categorical Perception* (ed. Harnad, S.) 161-188 (Cambridge Univ. Press, New York, 1987).
4. Werker, J. F. & Tees, R. C. *Infant Behav. Dev.* **7**, 49-63 (1984).
5. Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. *Science* **255**, 606-608 (1992).
6. Jusczyk, P. W. in *Measurement of Audition and Vision in the First Year of Postnatal Life: A Methodological Overview* (eds Gottlieb, G. & Krasnegor, N. A.) 195-222, (Ablex, Norwood, NJ, 1985).
7. Tucker, D. *Electroenceph. clin. Neurophysiol.* **87**, 154-163 (1993).
8. Woods, D. L. & Elmasian, R. *Electroenceph. clin. Neurophysiol.* **66**, 447-459 (1986).
9. Näätänen, R. & Gaillard, A. W. K. in *Tutorials in Event-related Potential Research: Endogenous Component* (eds Gaillard, A. W. K. & Ritter, W.) 119-142 (North-Holland, Amsterdam, 1983).
10. Näätänen, R. *Behav. Brain Sci.* **13**, 201-288 (1990).
11. Novak, G. P., Kurtzberg, D., Kreuzer, J. A. & Vaughan, H. G. Jr *Electroenceph. clin. Neurophysiol.* **73**, 295-305 (1989).
12. Scherg, M. & Berg, P. *BESA: Brain Electrical Source Analysis Handbook* (Max-Planck Institute for Psychiatry, Munich, 1990).
13. Witelson, S. F. & Pallie, W. *Brain* **96**, 641-646 (1973).
14. Chi, J. G., Dolling, E. C. & Gilles, F. H. *Ann. Neurol.* **1**, 86-93 (1977).
15. LeMay, M. in *Cerebral Dominance* (ed. Geschwind, N. & Galaburda, A. M.) 26-42 (Harvard Univ. Press, Cambridge, MA, 1984).
16. Bertoni, J. et al. *Brain Lang.* **37**, 591-605 (1989).
17. Vargha-Khadem, F. & Corballis, M. *Brain Lang.* **8**, 1-9 (1979).
18. Best, C. T., Hoffman, H. & Gianville, B. B. *Percept. Psychophys.* **31**, 75-85 (1982).
19. Molfese, D. L. & Burger-Judisch, L. M. in *Cerebral Lateralization: Theory and Research* (ed. Kitterle, F. L.) 71-102 (Erlbaum, Hillsdale, NJ, 1991).
20. Courchesne, E. in *Tutorials in Event-related Potential Research: Endogenous Component* (eds Gaillard, A. W. K. & Ritter, W.) 329-344 (North-Holland, Amsterdam, 1983).
21. Kurtzberg, D., Stapells, D. R. & Wallace, I. F. in *Identification of Infants with Developmental Disabilities* (eds Vietze, P. M. & Vaughan, H. G.) 160-180 (Allyn & Bacon, New York, 1988).
22. Mills, D. L., Coffey-Corina, S. A. & Neville, H. J. *J. Cogn. Neurosci.* **5**, 317-334 (1993).
23. Nelson, C. A. & deRegnier, R. *Dev. Neuropsychol.* **8**, 119-134 (1992).
24. Rothbart, M. K., Posner, M. I. & Boylan, A. in *The Development of Attention: Research and Theory* (ed. Enns, J. T.) 47-66 (Elsevier North-Holland, Amsterdam 1990).

25. Perrin, F., Pernier, J., Bertrand, D. & Echallier, J. F. *Electroenceph. clin. Neurophysiol.* **72**, 184–187 (1989).

ACKNOWLEDGEMENTS. We thank P. Compton, E. Gorr, D. Houston, L. McDougall, M. Murias, D. Perdue, T. Sumera and L. Thomas-Thrapp for technical help and M. Posner and D. Tucker for discussions. This study was supported in part by INSERM (G.D., S.D.), CNRS (S.D.) the NSF (S.D.), the Fyssen Foundation (G.D.) and the Société de Médecine Périnatale (G.D.) and by grants from the W.H. Keck Foundation and by the James S. McDonnell Foundation and Pew Memorial Trust support for the Center for the Cognitive Neuroscience of Attention.

Prion protein is necessary for normal synaptic function

John Collinge^{*†‡}, Miles A. Whittington[†],
Katie C. L. Sidle^{*}, Corinne J. Smith[§],
Mark S. Palmer^{*}, Anthony R. Clarke[§]
& John G. R. Jefferys[†]

^{*} Prion Disease Group, Department of Biochemistry and Molecular Genetics and [†] Department of Physiology and Biophysics, St Mary's Hospital Medical School, Imperial College, London W2 1PG, UK

[‡] Department of Neurology, St Mary's Hospital, London W2 1NY, UK

[§] Department of Biochemistry, University of Bristol, School of Medical Sciences, Bristol BS8 1TD, UK

THE prion diseases are neurodegenerative conditions, transmissible by inoculation, and in some cases inherited as an autosomal dominant disorder. They include Creutzfeldt–Jakob disease in humans and scrapie and bovine spongiform encephalopathy in animals. The prion consists principally of a post-translationally modified form of a host-encoded glycoprotein (PrP^C), designated PrP^{Sc} (ref. 1); the normal cellular function of PrP^C is, however, unknown. Although PrP is highly conserved among mammals and widely expressed in early embryogenesis, mice homozygous for disrupted PrP genes appear developmentally and behaviourally normal². PrP is a protein anchored to the neuronal surface by glycosylphosphatidylinositol, suggesting a role in cell signalling or adhesion. Here we report that hippocampal slices from PrP null mice have weakened GABA_A (γ -aminobutyric acid type A) receptor-mediated fast inhibition and impaired long-term potentiation. This impaired synaptic inhibition may be involved in the epileptiform activity seen in Creutzfeldt–Jakob disease and we argue that loss of function of PrP^C may contribute to the early synaptic loss³ and neuronal degeneration seen in these diseases.

We recorded electrophysiological responses from CA1 in hippocampal slices *in vitro*; there is a high level of PrP expression in the hippocampus⁴. Responses from PrP null slices occasionally contained multiple population spikes (Fig. 1a) and pairs of intracellular action potentials (Fig. 1b), which were never seen in controls. These occurred reliably after trains of stimuli that weaken inhibition⁵. The genetic background of the PrP null mice is derived from both 129/Sv and C57BL/6J; we therefore used C57BL/6J, 129/Sv and an F₁ cross between these as controls. There were no differences between the two inbred lines and the F₁ animals with respect to the findings reported here. Responses evoked in CA1 pyramidal cells normally terminate with prominent fast and slow inhibitory postsynaptic potentials (i.p.s.ps). Fast i.p.s.ps appeared weaker than normal in the PrP null slices (see below). Excitatory postsynaptic potentials (e.p.s.ps) were qualitatively normal. We concluded that the lack of PrP had no effect on the passive properties of CA1 neurons (Fig. 1 legend).

Single-electrode voltage clamp was used to investigate fast i.p.s.ps in slices in which e.p.s.ps and slow i.p.s.ps were blocked pharmacologically (Fig. 2)⁶. Stimulation near the recording site reliably excited inhibitory interneurons directly, and elicited monosynaptic inhibitory postsynaptic currents (i.p.s.cs) in all

pyramidal cells. The time constant of the decay phase of maximal i.p.s.cs did not differ between groups but the rising phase was significantly slower in PrP null slices (Fig. 2b). Maximal monosynaptic i.p.s.cs evoked at a series of holding potentials yielded current–voltage curves (Fig. 2c). The reversal potentials for fast i.p.s.cs in PrP null slices were significantly depolarized compared with controls (-62.2 compared with -69.2 mV). The fast i.p.s.c. conductance, estimated from individual current–voltage curves, was significantly lower in PrP null slices (57 nS) than in any control group (means of 86–101 nS). Differences in GABA_A-mediated fast i.p.s.ps are at least partly postsynaptic, because the initial response to exogenous GABA showed a similar depol-

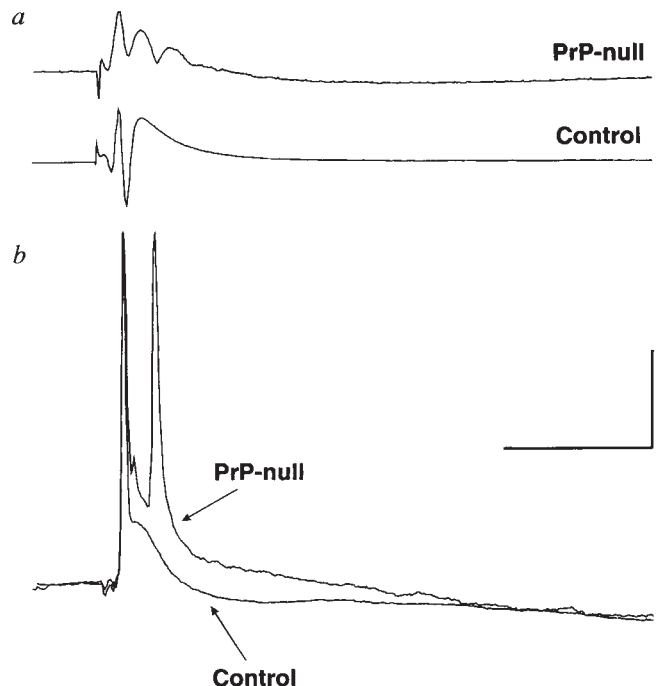


FIG. 1 a, Extracellular responses evoked by afferent stimulation and recorded from the CA1 pyramidal cell layer in hippocampal slices from PrP null and from 129/Sv $\times$ C57BL/6J (control) mice reveal extra population spikes in the former. This occurred in a minority of slices from these animals without any special treatment, and reliably after repetitive stimulation (for example, 100 stimuli in 1 s, repeated every 30 s for 2.5 min). b, Intracellular recordings from the same slices reveal a pair of action potentials and the lack of a hyperpolarizing fast i.p.s.p. elicited in the CA1 pyramidal layer cell from the PrP null slice. There were no differences in resting potential between them and the three control groups ($F_{3,39} = 2.38$, $P = 0.08$, 1-way ANOVA; mean for PrP null slices 65 ± 1 mV, and for combined controls 66 ± 1 mV). However, the C57BL/6J mice had higher input resistances than any of the other groups (64.5 ± 4.9 M Ω compared with means of 40.0–49.2 M Ω for the other groups; $F_{3,39} = 8.11$, $P < 0.001$, ANOVA; C57BL/6J differed from each other group, $P < 0.05$; Bonferroni multiple comparisons). However, this difference cannot be attributed to the presence or absence of PrP as PrP null slices did not differ from 129/Sv or 129/Sv $\times$ C57BL/6J. C57BL mice have different neuronal membrane composition from many other mouse strains²². Scale bars: a, 5 mV; b, 20 mV; a and b, 20 ms.

METHODS. Genotype of PrP null mice was confirmed by polymerase chain reaction (PCR) analysis of tail DNA². Adult male mice were killed by cervical dislocation and transverse 400- μ m slices of dorsal hippocampus were prepared and maintained at 35 °C and perfused with ACSF (NaCl 135 mM, NaHCO₃ 16 mM, CaCl₂ 2 mM, KCl 3 mM, NaH₂PO₄ 1.25 mM, MgCl₂ 1 mM, glucose 10 mM) under warm, wet, 95% O₂, 5% CO₂. After >30 min recovery, recordings were made with glass micropipettes from the CA1 pyramidal layer (extracellular electrodes were filled with 2 M NaCl, resistances 5–15 M Ω , and intracellular with 4 M potassium acetate, resistances 35–65 M Ω)²³. Bipolar stimulating electrodes were placed in stratum radiatum. The recordings were made blind to mouse genotype.

|| To whom correspondence should be addressed.

FIG. 2 Pharmacologically isolated monosynaptic fast i.p.s.cs recorded from CA1. **a**, Averages of 8 consecutive traces for PrP null slices and controls (129/Sv $\times$ C57BL/6J F_1 crosses) were recorded as outward currents at a holding potential of -40 mV and are superimposed to compare their amplitudes and shapes, which were close to the modal values of their respective groups. **b**, i.p.s.cs from PrP null slices were slower. The decay-phase time constant did not differ between groups ($F_{3,34} = 3.26$, $P = 0.125$, 1-way ANOVA), but the rising phase was significantly slower in PrP null slices (3.33 ± 0.35 ms) than in C57BL/6J and 129/Sv slices (2.36 ± 0.15 and 2.44 ± 0.13 ms respectively; $F_{3,28} = 3.26$, $P = 0.038$; Bonferroni multiple comparisons at $P < 0.05$). **c**, i.p.s.c. amplitudes were measured at holding potentials between -100 and -40 mV. The lower amplitude of the PrP null trace in **a** corresponds to a decrease in the slope of the I - V curve constructed from means of 14 PrP null ($\bullet$) and 9 control ($\circ$) cells. i.p.s.c. conductances estimated from the curves for individual experiments showed that those for the PrP null group were significantly lower than controls ($F_{3,33} = 8.62$, $P < 0.001$, and Bonferroni multiple comparisons at $P < 0.05$). The reversal potential in the PrP null group was significantly more depolarized ($H_3 = 8.95$, $P < 0.03$; Kruskal-Wallis 1-way ANOVA on ranks; PrP null group differed from C57BL/6J, $P < 0.05$; Dunn's test). Slow i.p.s.ps mediated by GABA_B receptors isolated pharmacologically did not differ between the groups (PrP null group conductance = 18 ± 1 nS, reversal potential = 90 ± 4 mV; controls 19 ± 2 nS and 88 ± 3 mV respectively). **d**, The change in GABA_A-mediated i.p.s.cs was at least partly postsynaptic because responses to exogenous GABA showed similar shifts in reversal potential of the initial component, shown here at -60 mV. **METHODS**. Monosynaptic fast, bicuculline-sensitive i.p.s.cs were evoked in CA1 by proximal stimulation, using drugs ($20 \mu\text{M}$ 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX), $50 \mu\text{M}$ 2-amino-5-phosphonovaleate (APV), $200 \mu\text{M}$ 2-OH saclofen) to block e.p.s.ps and slow i.p.s.ps⁶. Intracellular recordings were made with sharp micropipettes containing 50 mM QX314 to improve clamp efficiency. Recordings were made using single-electrode voltage clamp (Axoclamp); clamp efficiency was kept constant between the experimental groups (in the range 72–80%). Stimuli were applied close to stratum pyramidale, $\sim 50 \mu\text{m}$ from the recording site. In each case the stimulus required for maximal response was determined and kept constant for the rest of the experiment. Responses of these receptors to exogenous GABA were recorded under similar conditions, using pressure ejection of GABA (1 mM) from a pipette positioned in stratum pyramidale to give maximal response (pressure pulse of 1.2 kg cm^{-2} for 100 ms). Monosynaptic, slow, 2-OH-saclofen-sensitive i.p.s.ps were evoked similarly to the fast i.p.s.cs, but in the presence of $20 \mu\text{M}$ CNQX, $50 \mu\text{M}$ APV and $30 \mu\text{M}$ bicuculline methiodide. (Scale bars: **a**, 2 nA, 20 ms; **d**, 2 nA, 2 s.)

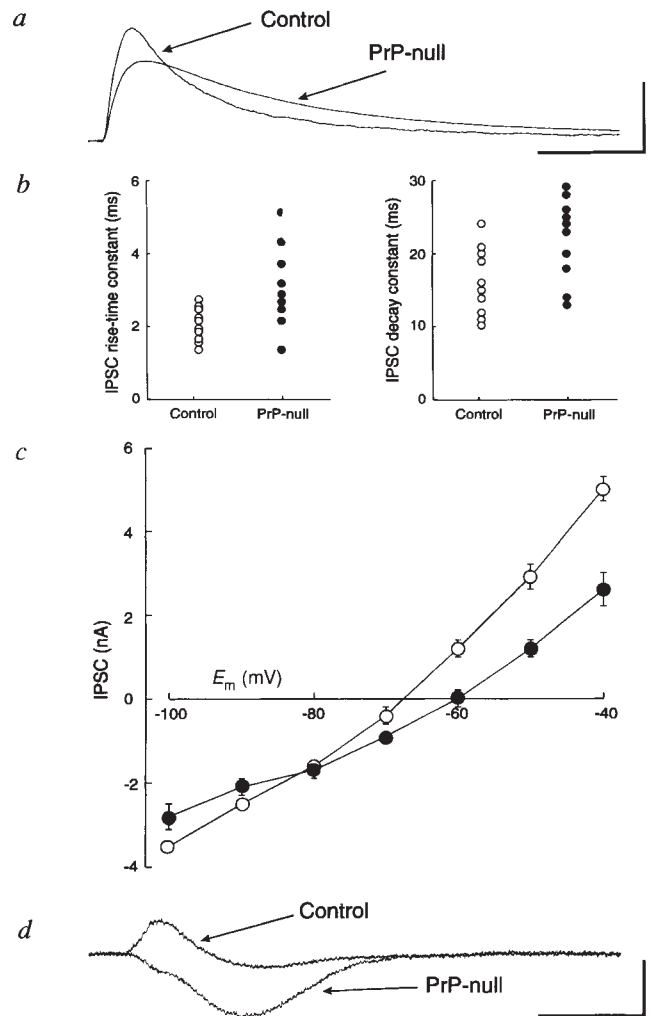
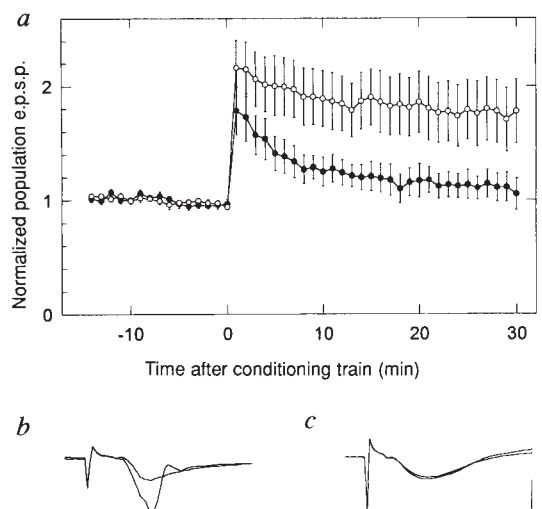


FIG. 3 LTP in the CA1 region of hippocampal slices is significantly weaker in PrP null mice ($\bullet$) than in control animals (C57BL $\times$ 129/Sv) ($\circ$) whether in the presence or in the absence (**a**) of synaptic inhibition mediated by GABA_A receptors. Pairs of traces from control (**b**) and PrP null (**c**) slices incubated in $30 \mu\text{M}$ bicuculline to block fast i.p.s.ps show superimposed responses just before and 30 min after the conditioning train. LTP was measured 21–30 min after conditioning as a fraction of the mean pre-conditioning value. With inhibition blocked, LTP reached 1.77 ± 0.07 ($n = 12$) in controls and 1.12 ± 0.06 ($n = 13$) in PrP null slices (LTP was stable over this period, $F_{9,220} = 0.04$, $P > 0.99$, 2-way ANOVA, and the experimental groups were significantly different, $F_{1,220} = 51.7$, $P < 0.001$). Similar results were obtained with intact inhibition; in PrP null mice mean potentiation after 21–30 min was 1.21 ± 0.05 compared with 1.59 ± 0.05 in controls (LTP was stable over this period, $F_{9,80} = 0.01$, $P > 0.99$; PrP null mice differed significantly from controls, $F_{1,80} = 21.3$, $P < 0.001$). Weaker LTP in PrP null mice could not be due to impairment of the NMDA component of e.p.s.ps because there was no difference when isolated pharmacologically; amplitudes in PrP null and control mice were, respectively, 1.7 ± 0.4 and 2.1 ± 0.1 mV, at a resting potential of -70 mV, and of 7.8 ± 2.0 and 8.0 ± 1.5 mV at -40 mV ($n = 5$ in each case). **METHODS**. Extracellular recordings were from stratum radiatum. Compound synaptic potentials were evoked by electrical stimulation at a site 0.3 – 0.5 mm away in the same layer every 15 s. Groups of 4 potentials were averaged to yield one measurement per min, of the slope of the rising phase of the negative population e.p.s.p. The conditioning train of 100 stimuli at 100 Hz was given after 15 min of stable baseline. The initial values of e.p.s.ps were adjusted to about half-maximal. Experiments in the presence of $30 \mu\text{M}$ bicuculline methiodide, to



remove potentially confounding changes in synaptic inhibition, used slices with CA3 removed to avoid epileptic activity. Finally, the NMDA component of the e.p.s.p. evoked by maximal stimulation was measured intracellularly in the presence of $20 \mu\text{M}$ CNQX, $30 \mu\text{M}$ bicuculline and $200 \mu\text{M}$ 2-OH-saclofen. (Scale bars: 2 mV, 20 ms.)

arized reversal (Fig. 2d). There is also some selectivity because GABA_B-mediated slow i.p.s.ps were unaffected (Fig. 2 legend).

We investigated whether long-term potentiation (LTP), a form of synaptic plasticity associated with learning and memory⁷, was preserved. The slope of the rising phase of evoked potentials estimated e.p.s.p. strength across the neuronal population. The conditioning stimulus train (1 s, 100 Hz at time 0; Fig. 3) reliably induced transient post-tetanic potentiation. It also induced LTP, lasting more than 60 min, in control slices. However, PrP null slices revealed significantly weaker LTP in both the absence and the presence of bicuculline (Fig. 3). The NMDA (*N*-methyl-D-aspartate) receptor is necessary for LTP in CA1, but there was no difference in the pharmacologically isolated NMDA component of maximal evoked e.p.s.ps in PrP null and control mice (Fig. 3 legend).

We found several phenotypic differences between PrP null mice and the inbred lines from which they were derived. The tendency to multiple spikes in the CA1 region can be attributed to impaired fast i.p.s.ps/i.p.s.cs in cells from PrP null mice. Smaller i.p.s.p./c. amplitudes could arise at many locations within inhibitory synapses. However, postsynaptic mechanisms are implicated in the depolarized reversal potentials of fast i.p.s.ps and of responses to exogenous GABA, probably abnormal Cl⁻ gradients or selectivity of the GABA-gated ion channels. Although the mechanisms of depolarizing GABA_A responses remain unclear⁸, they may be mediated by extrasynaptic receptors⁹⁻¹¹. In the present study, i.p.s.cs had significantly slower rising phases, consistent with a longer diffusion path.

LTP in CA1 of hippocampal slices is significantly weaker in PrP null mice. The PrP null mice used here were reported to perform the Morris swim task as well as controls⁵. This is surprising given that this behavioural task depends on hippocampal integrity, and that LTP has been linked with learning⁷. However, there are precedents for discrepancies between LTP and hippocampal learning tasks¹². The most plausible mechanism for the poor LTP in PrP null mice is that the weakened inhibition enhances *in vivo* the NMDA component of normal, low-frequency e.p.s.ps¹³, and that this abnormal activation of NMDA receptors impairs LTP induction^{14,15}.

The i.p.s.c. abnormalities in PrP null mice are consistent with GABA having to diffuse further, to a significant proportion of receptors with extrasynaptic properties⁹. This could arise if the receptors were not properly localized, or if synaptic cleft architecture had changed. It will be important to study the location and topology of PrP in neuronal terminals and to examine synaptic ultrastructure in PrP null mice. Although prion diseases are neurodegenerative conditions, with little evidence for disease outside the nervous system, PrP is expressed in other tissues¹⁶; it remains to be established if PrP has a functional role in these locations.

The current model of prion propagation proposes that PrP^{Sc}, either inoculated, or generated as a stochastic event from mutant PrP^C (or more rarely from wild-type PrP^C), interacts with PrP^C and promotes its conversion to PrP^{Sc} (refs 17-19). PrP^{Sc} accumulates, is partially protease-resistant and forms amyloid. The report that PrP null mice develop and behave normally² focused attention on PrP^{Sc} accumulation as the cause of neuronal damage. However, prion dementia can occur without the cytoarchitectural changes or amyloid deposition characteristic of these diseases²⁰. The abnormalities of synaptic inhibition reported here are relevant to the epileptiform discharges seen in Creutzfeldt-Jakob disease and scrapie-infected mice²¹. While inherited prion diseases are dominant disorders, suggesting a gain rather than loss of function, disease may result from a dominant negative effect, with PrP^{Sc}, generated from mutant PrP, leading to progressive functional loss of PrP^C. Loss of PrP^C in an adult nervous system may have more severe effects than in a developing one, where adaptation may occur. Our findings reinstate the possibility that prion neurodegeneration may be due, at least in part, to loss of function of PrP. □

Received 12 April; accepted 30 June 1994.

1. Prusiner, S. B. *Science* **252**, 1515-1522 (1991).
2. Bueler, H. et al. *Nature* **356**, 577-582 (1992).
3. Clinton, J., Forsyth, C., Royston, M. C. & Roberts, G. W. *Neuroreport* **4**, 65-68 (1993).
4. DeArmond, S. J. et al. *Neurology* **37**, 1271-1280 (1987).
5. Thompson, S. M. & Gahwiler, B. H. *J. Neurophysiol.* **61**, 501-511 (1989).
6. Davies, C. H., Davies, S. N. & Collingridge, G. L. *J. Physiol., Lond.* **424**, 513-531 (1990).
7. Bliss, T. V. P. & Collingridge, G. L. *Nature* **361**, 31-39 (1993).
8. Grover, L. M., Lambert, N. A., Schwartzkroin, P. A. & Teyler, T. J. *J. Neurophysiol.* **69**, 1541-1555 (1993).
9. Alger, B. E. & Nicoll, R. A. *J. Physiol., Lond.* **328**, 125-141 (1982).
10. Brown, D. A., Higgins, A. J., Marsh, S. & Smart, T. G. *Adv. Biochem. Psychopharmac.* **29**, 321-326 (1981).
11. Allan, R. D., Evans, R. H. & Johnston, G. A. *Br. J. Pharmac.* **70**, 609-615 (1980).
12. Brace, H. M., Jefferys, J. G. R. & Mellanby, J. J. *J. Physiol., Lond.* **368**, 343-357 (1985).
13. Herron, C. E., Williamson, R. & Collingridge, G. L. *Neurosci. Lett.* **61**, 255-260 (1985).
14. Coan, E. J., Irving, A. J. & Collingridge, G. L. *Neurosci. Lett.* **105**, 205-210 (1989).
15. Huang, Y. Y., Colino, A., Selig, D. K. & Malenka, R. C. *Science* **255**, 730-733 (1992).
16. Bendheim, P. E. et al. *Neurology* **42**, 149-156 (1992).
17. Prusiner, S. B. et al. *Cell* **63**, 673-686 (1990).
18. Weissmann, C. *Nature* **349**, 569-571 (1991).
19. Palmer, M. S., Dryden, A. J., Hughes, J. T. & Collinge, J. *Nature* **352**, 340-342 (1991).
20. Collinge, J. et al. *Lancet* **336**, 7-9 (1990).
21. Jefferys, J. G. R., Empson, R. M., Whittington, M. A. & Prusiner, S. B. *Neurobiology of Disease* **1**, 3-15 (1994).
22. Smith, T. L. *J. Pharmac. exp. Ther.* **232**, 702-707 (1985).
23. Empson, R. M. & Jefferys, J. G. R. *J. Physiol., Lond.* **465**, 595-614 (1993).

ACKNOWLEDGEMENTS. This work was funded by the Wellcome Trust, BBSRC Biology of Spongiform Encephalopathies Programme and the David and Frederick Barclay Foundation. We thank S.B. Prusiner and C. Weissmann for supplying PrP gene ablated mice.

Inhibition of the cardiac protein kinase A-dependent chloride conductance by endothelin-1

Andrew F. James*†, Lai-Hua Xie†, Yasushi Fujitani*, Seiji Hayashi† & Minoru Horie†§

* International Research Laboratories,

Ciba-Geigy Japan Ltd, Takarazuka 665, Japan

† Department of Cellular and Molecular Physiology,

National Institute for Physiological Sciences, Okazaki 444, Japan

‡ 3rd Department of Internal Medicine, Faculty of Medicine, Kyoto University, Kyoto 606, Japan

ENDOTHELIN-1 is a peptide hormone constitutively secreted by vascular and endocardial endothelial cells¹⁻³. Secretion of endothelin-1 is increased under certain pathophysiological conditions, including coronary vasospasm, cardiac ischaemia and myocardial infarction. We have examined the effect of endothelin-1 on the protein kinase A (PKA)-dependent chloride current in voltage-clamped guinea pig ventricular myocytes. This conductance, induced by catecholamines through β -adrenergic receptors, counteracts the simultaneously increased L-type calcium current by shortening the action potential duration⁴⁻⁶. We report here that endothelin-1, acting through ET_A (endothelin-1-selective) receptors, inhibited the current through a pertussis toxin-sensitive mechanism, analogous to muscarinic receptors, by reducing the intracellular cyclic AMP concentration. This effect of endothelin-1 should help protect the ventricle against potentially arrhythmogenic shortening of the action potential during ischaemia when the circulating levels of catecholamines are increased.

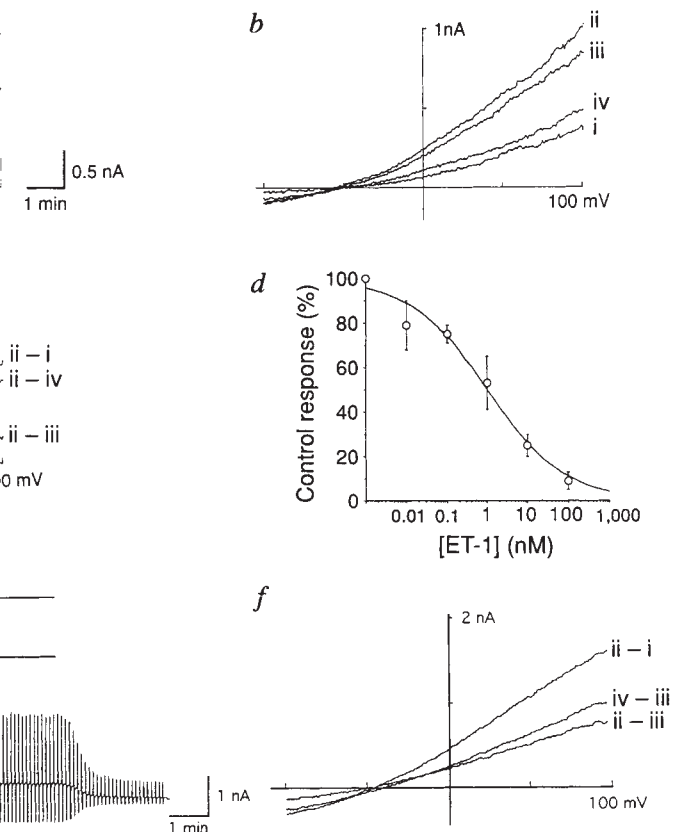
Exposure of a voltage-clamped ventricular myocyte dialysed internally with 20 mM Cl⁻ pipette solution and superfused with 150 mM Cl⁻ external perfusion solution (EPS) to 100 nM of the β -adrenergic agonist isoprenaline induced outwardly rectifying currents typical of the cardiac cAMP-dependent kinase (PKA)-dependent Cl⁻ conductance⁴⁻⁶ (Fig. 1). The myocyte was held at 0 mV and voltage ramps were applied to obtain the instantaneous current-voltage relationship. The resting current reversed at -30 mV, close to the estimated equilibrium potential for Cl⁻

§ To whom correspondence should be addressed.

FIG. 1 Endothelin-1 inhibits in a concentration-dependent manner the isoprenaline-induced current. **a**, Chart recording of the whole-cell current. Holding potential = 0 mV. The vertical lines represent the current due to 1-s voltage ramps from +100 mV to -100 mV applied every 6 s using a signal generator. Bars mark exposure to extracellular 100 nM isoprenaline (Iso) and 0.1 nM or 10 nM endothelin-1 (ET-1). **b**, Current-voltage relationships as indicated in **a**. **i**, The resting current, after exposure to **ii**, 100 nM Iso, **iii**, 0.1 nM and **iv**, 10 nM ET-1. **c**, Difference currents showing the Iso-induced current (**ii** - **i**), the current inhibited by 0.1 nM (**ii** - **iii**) and 10 nM (**ii** - **iv**) ET-1. **d**, The ET-1 concentration-dependence of the inhibition of the 100 nM Iso-induced current at +100 mV (737.3 ± 63.7 pA). Points represent mean $\pm$ s.e. ($n \geq 3$). The curve represents a fit to a logistic equation. Hill coefficient = 0.45 ± 0.06 . **e**, Chart recording from another cell. Bars indicate exposure to 30 nM and 1 μ M Iso and 3 nM ET-1. **f**, Difference currents in **e** showing the 30 nM Iso-induced current (**ii** - **i**), the 3 nM ET-1 inhibited current (**ii** - **iii**) and the current restored on increasing [Iso] to 1 μ M in the continued presence of 3 nM ET-1 (**iv** - **iii**).

METHODS. Experiments were done at $\sim 36^\circ\text{C}$ using guinea-pig left ventricle myocytes as described previously²⁶ with certain exceptions. Cells were stored in RPMI medium containing Ca^{2+} ($\sim 21^\circ\text{C}$). Pipettes, filled with the 20 mM Cl^- solution before experiments, had a tip resistance of 3–5 M Ω . The pipette solution contained BAPTA and sodium creatine phosphate in place of EGTA and Tris creatine phosphate. 100 μ M GTP

ions (Fig. 1*bi*). Exposure to isoprenaline resulted in about a threefold increase in the membrane conductance and holding current without altering the reversal potential (Fig. 1*a, bi, ii*). Subsequent exposure to 0.1 nM or 10 nM endothelin-1 reduced the current in a concentration-dependent manner; 0.1 nM endothelin-1 reduced the isoprenaline induced current (Fig. 1*c, ii - i*) by $\sim 20\%$ (Fig. 1*c, ii - iii*) and 10 nM endothelin-1 reduced the current by $\sim 80\%$ (Fig. 1*c, ii - iv*). Endothelin-1 did not affect the reversal potential, demonstrating that endothelin-1 reduced

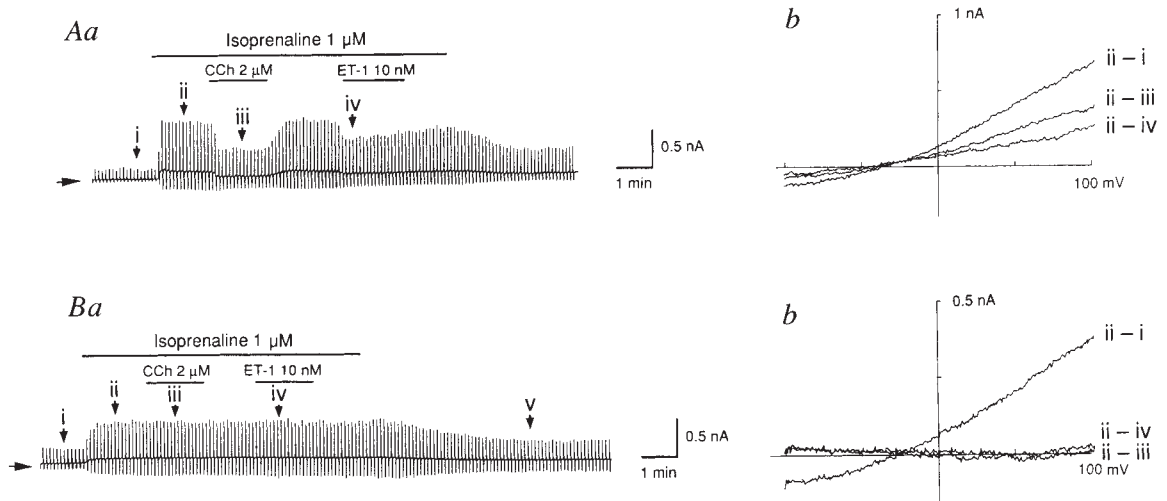


was added to the pipette solution to reduce run-down of G-protein-mediated responses²⁶. The Ca^{2+} - K^+ -free extracellular perfusion solution (EPS) contained 150 mM NaCl. Liquid junction potentials were corrected using the voltage-offset on the patch-clamp amplifier. Currents were recorded to a chart recorder and simultaneously digitized (1 kHz) and stored on-line for later computer analysis.

the membrane Cl^- conductance specifically (Fig. 1*b, c*). It also inhibited in a concentration-dependent manner the isoprenaline-induced currents with a half-maximally effective concentration of $1.0 \text{ nM} \pm 0.3 \text{ nM}$. Exposure of myocytes to a submaximal (30 nM) concentration of isoprenaline induced a Cl^- current which was inhibited to $58.7\% \pm 3.7$ ($n = 3$) by 3 nM endothelin-1 (for example, Fig. 1*e, f*). Subsequent elevation of isoprenaline to 1 μ M in the continued presence of endothelin-1 completely overcame the inhibitory action, resulting in a current

FIG. 2 The effect of ET-1 is pertussis toxin (PTX) sensitive. **A**, Control cells treated with buffer alone. **a**, Chart record of whole-cell current. Bars indicate exposure to 1 μ M Iso, 2 μ M carbachol (CCh) and 10 nM ET-1. **b**, Difference currents as indicated in **a** showing the Iso-induced current (**ii** - **i**), the CCh-sensitive current (**ii** - **iii**), and the ET-1 sensitive current (**ii** - **iv**). **B**, PTX-treated cells. **a**, Bars as in **A**. **b**, Difference currents as indicated in **a**.

METHODS. As described in Fig. 1 except that cells were subsequently incubated for a further 90 min at 36°C with 7.5 ng ml^{-1} PTX or buffer alone.



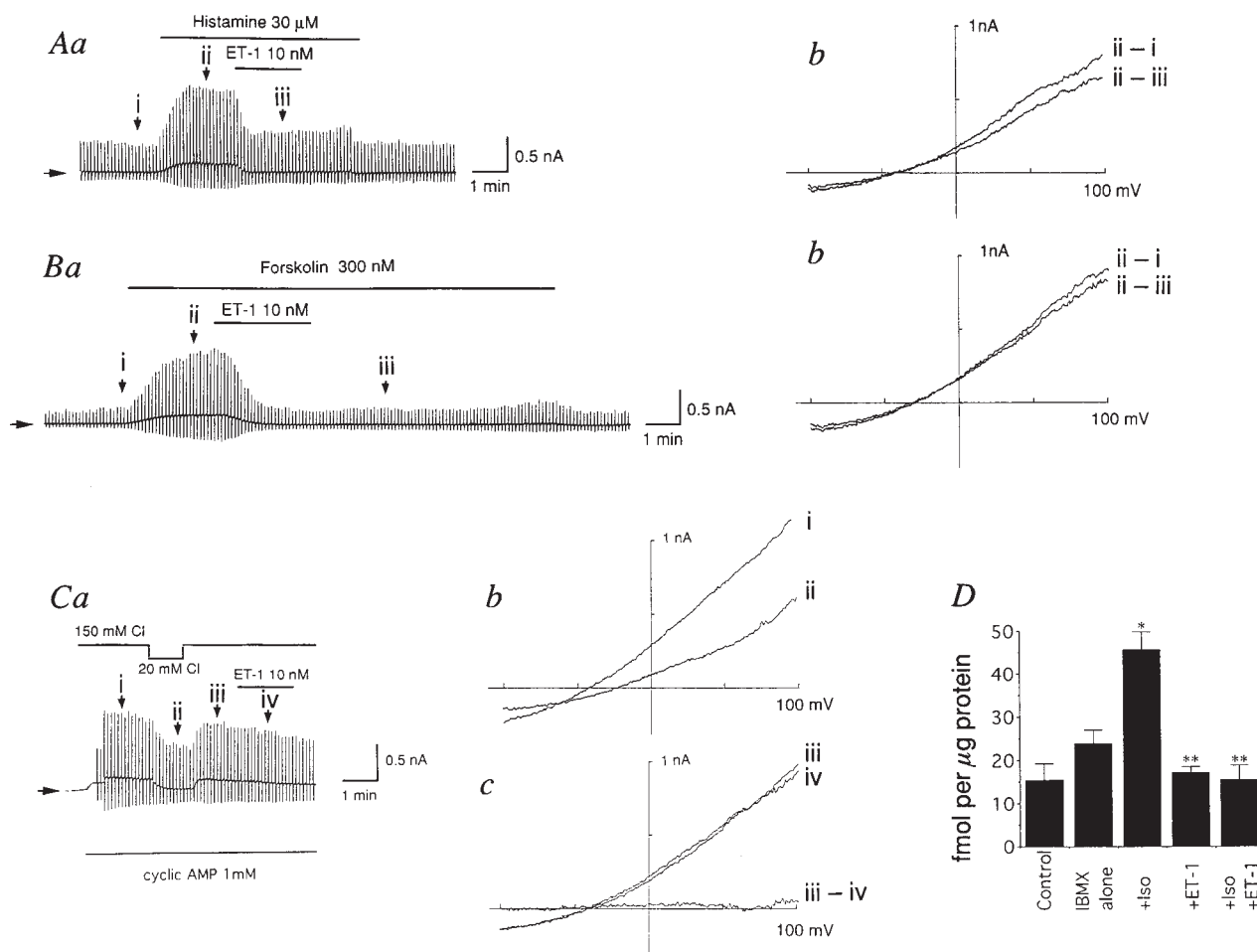


FIG. 3 ET-1 inhibits the current by reducing the intracellular cyclic AMP level. **A**, Inhibition of the histamine-induced current. **a**, Chart record. Bars indicate exposure to 30 μ M histamine and 10 nM ET-1. **b**, Difference currents as indicated in **a** showing the histamine-induced current (ii - i) and ET-1-sensitive current (ii - iii). **B**, Inhibition of the forskolin-induced current. **a**, Chart record. Bars indicate exposure to 300 nM forskolin and 10 nM ET-1. **b**, Difference currents as indicated in **a** showing the forskolin-induced current (ii - i) and the ET-1-sensitive current (ii - iii). **C**, Absence of effect of ET-1 on the current induced by internal dialysis with cAMP. **a**, Chart record. Bar at the bottom of trace indicates internal dialysis with 1 mM AMP immediately upon whole-cell formation. Bars at the top of the trace indicate change of the external Cl^- concentration to 20 mM and exposure to 10 nM ET-1. **b**, Dependency of current-voltage relationship on external Cl^- ions. In the presence of 150 mM (i) and 20 mM (ii) $[\text{Cl}^-]_o$. **c**, Current-voltage relationships show-

ing the control current (iii) and after exposure to 10 nM ET-1 (iv) and the ET-1-sensitive current (iii - iv). **D**, ET-1 reduces the ventricular cAMP level. Coordinate axis of the histogram shows the ventricular cAMP level (fmol cAMP per μ g protein) after exposure to Tyrode's alone, 0.5 mM isobutylmethyl xanthine (IBMX), 200 nM Iso + IBMX, 10 nM ET-1 + IBMX and Iso + ET-1 + IBMX. *, $P < 0.05$ compared with cAMP content measured in the presence of IBMX alone using Student's *t*-test. **, $P < 0.05$ compared with cAMP content measured in the presence of Iso.

METHODS. A-C, As described in Fig. 1. **C**, 20 mM EPS was made by substitution of 130 mM Cl^- with aspartate. **D**, Guinea-pig left ventricular slices were incubated for 3 min at 37 $^{\circ}\text{C}$ in Tyrode's supplemented with the compounds as indicated. The cAMP content was determined by radioimmunoassay and the protein content by BCA assay.

134.0% $\pm$ 22.5% of the 30 nM isoprenaline-induced current. This result suggests competition between isoprenaline and endothelin-1 in the activation of the Cl^- current.

The cardiac PKA-dependent Cl^- current reflects the underlying adenylate cyclase activity modulated by the competitive and antagonistic actions of G_s and G_i proteins^{6,7}. Because endothelin receptors belong to the G-protein-coupled receptor superfamily^{1,2}, the effect of pertussis toxin (PTX) was investigated. Although both carbachol and endothelin-1 inhibited the isoprenaline-induced current in control cells (Fig. 2A), their effects were abolished in cells treated with PTX (Fig. 2B). Histamine and forskolin, acting through adenylate cyclase^{7,8}, both induced currents which were inhibited by endothelin-1 at concentrations similar to those effective against the isoprenaline-induced current (Fig. 3A, B). When the cell was dialysed internally with cAMP, a large Cl^- -selective outwardly rectifying current was induced soon after formation of the whole cell configuration (Fig. 3C). This current was not

inhibited by endothelin-1. Thus, endothelin-1 acted by reducing the intracellular cAMP concentration through inhibition of adenylate cyclase.

Endothelin-1 inhibited the isoprenaline-induced increase in the cAMP content of ventricular slices measured by radioimmunoassay (Fig. 3D). It effectively abolished the increase in cAMP induced by 200 nM isoprenaline (Fig. 3D) at a concentration (10 nM) that, on average, inhibited the Cl^- current by only 80% (Fig. 1d). The efficacy of the endothelin receptors may be reduced by enzymatic treatment before electrophysiological experiments. Variation was seen in both the degree of responsiveness to isoprenaline (for example, Fig. 1a, e) and to endothelin-1 (Fig. 1d) and seemed to depend largely on the cell isolation, the enzyme used and the duration of digestion. Certainly, 10 nM endothelin-1 was able to inhibit the current almost completely in some experiments (Fig. 3B).

Competitive binding assays using radiolabelled endothelin-1 and the ET_A -selective antagonist, BQ-123 (ref. 9), and the ET_B -

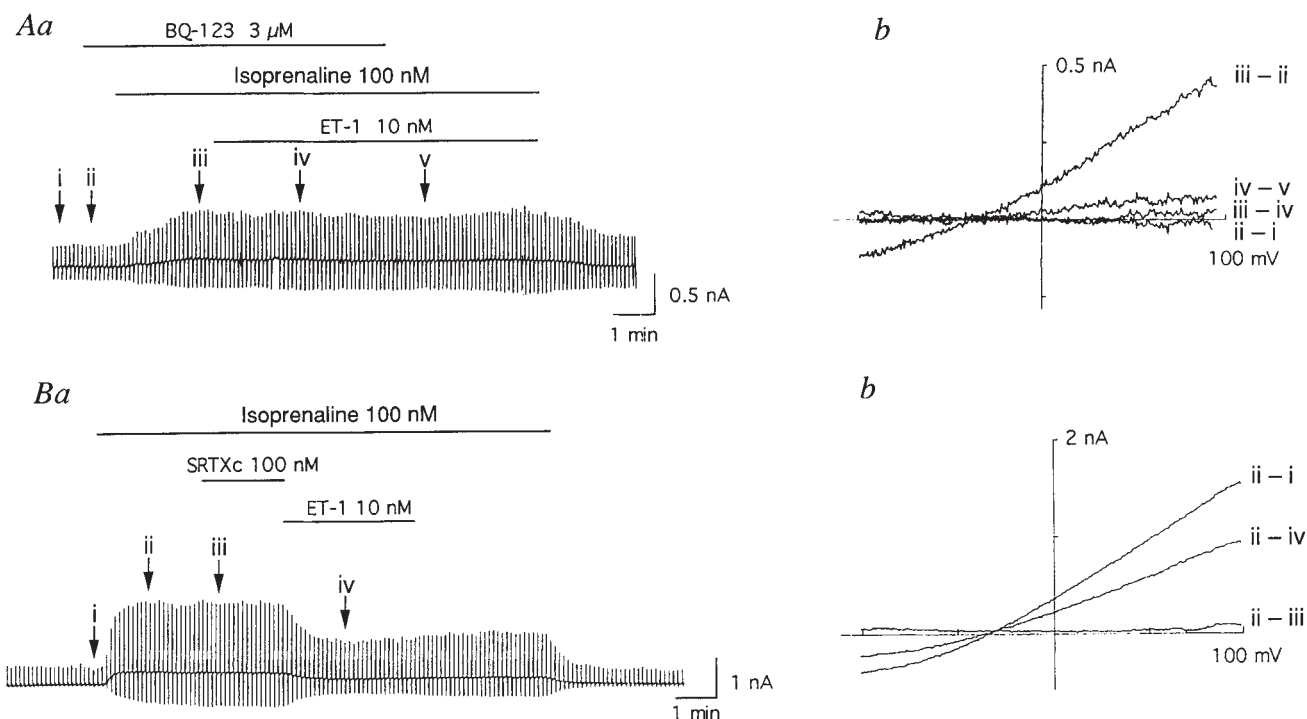


FIG. 4 Sensitivity of the response to endothelin receptor-selective compounds. **A**, Abolition of the inhibition by ET-1 by pre-incubation with 3 μ M BQ-123. **a**, Chart record of the whole-cell current. Bars indicate exposure to 3 μ M BQ-123, 100 nM Iso, and 10 nM ET-1. **b**, Difference currents as indicated in **a** showing the effect of BQ-123 on the resting current (ii - i), the Iso-induced current (iii - ii), the ET-1-sensitive current

(iii - iv), and the ET-1-sensitive current after washout of BQ-123 (iv - v). **B**, Absence of effect of 100 nM sarafotoxin S6c (SRTXc) on the Iso-induced current. **a**, Chart record of the whole-cell current. Bars indicate exposure to 100 nM Iso, 100 nM SRTXc and 10 nM ET-1. **b**, Difference currents as indicated in **a** showing the Iso-induced currents (ii - i), the SRTXc- (ii - iii) and ET-1-sensitive currents (ii - iv).

selective agonist, IRL 1620 (refs 10, 11), revealed both ET_A (BQ-123, 78.9%; IRL 1620, 75.7%) and ET_B (BQ-123, 21.1%; IRL 1620, 24.3%) receptors^{1,2} in ventricular membranes (data not shown). The ET_A-selective antagonist BQ-123 completely prevented the inhibitory action of ET-1 (Fig. 4A). Moreover, the ET_B-selective agonist sarafotoxin S6c (SRTXc)¹² did not inhibit the isoprenaline-induced current, although subsequent exposure of the same cell to endothelin-1 resulted in marked reduction of the current (Fig. 4B). These data suggest strongly that the effect of endothelin-1 was mediated almost entirely by ET_A receptors.

Endothelin-1 has been shown to stimulate inositol polyphosphate (InsP) accumulation in rat^{13,14} and cat¹⁴ adult ventricular myocytes by a PTX-insensitive mechanism. Because the inhibition of the rat ventricular adenylate cyclase by endothelin-1 was abolished by PTX¹³, InsPs were not involved in the inhibitory action of endothelin-1. Our data are consistent with this conclusion; intracellular Ca²⁺ was chelated to a very low free concentration (0.1 nM–1 nM) with *O*, *O'*-bis(2-aminophenyl)-ethyleneglycol-*N,N,N',N'*-tetraacetic acid (BAPTA) and it is unlikely that there was significant G-protein-activated phospholipase C activity^{15,16}. Endothelin-1 has recently been shown to inhibit cAMP formation and stimulate InsP accumulation through ET_A receptors in the human atrium¹⁷.

Endothelin-1 has a positive inotropic effect¹⁸ whose mechanism remains controversial^{13,19–22}. In contrast, we have demonstrated an inhibitory action of endothelin-1 analogous to that of muscarinic agonists. Because it reduces the intracellular cAMP content, endothelin-1 can be expected to inhibit the catecholamine-induced increases in Ca²⁺ and delayed rectifier K⁺ current. We suggest that endothelin-1, secreted by coronary artery and endocardial endothelial cells, acts as a local paracrine hormone^{1,2} to modify the cardiac responses to catecholamines. Endothelins have previously been reported to inhibit the positive inotropic response to sympathetic stimulation²³ and to exert a negative chronotropic effect on rat atrial myocytes²⁴. A negative chronotropic effect of endothelin-1 mediated by ET_A receptors

through a pertussis-toxin G-protein/adenylate cyclase inhibition pathway similar to that described here has very recently been demonstrated in the guinea pig atrium²⁵. Inhibition of the Cl[−] conductance by endothelin-1 can be expected to counteract the potentially arrhythmogenic shortening of the action potential due to ATP-sensitive K⁺ channel activation during coronary ischaemia. □

Received 30 March; accepted 1 June 1994.

- Masaki, T., Yanagisawa, Y. & Goto, K. *Med. Res. Rev.* **12**, 391–421 (1992).
- Haynes, W. G. & Webb, D. J. *Clin. Sci.* **84**, 485–500 (1993).
- Mebazaa, A. et al. *Am. J. Physiol.* **265**, H1841–H1846 (1993).
- Bahinski, A., Nairn, A. C., Greengard, P. & Gadsby, D. C. *Nature* **340**, 718–721 (1989).
- Harvey, R. D. & Hume, J. R. *Science* **244**, 983–985 (1989).
- Harvey, R. D., Clark, C. D. & Hume, J. R. *J. gen. Physiol.* **95**, 1077–1102 (1990).
- Hwang, T.-C., Horie, M., Nairn, A. C. & Gadsby, D. C. *J. gen. Physiol.* **99**, 465–489 (1992).
- Harvey, R. D. & Hume, J. R. *J. cardiovasc. Electrophysiol.* **1**, 309–317 (1990).
- Ihara, M. et al. *Life Sci.* **50**, 247–255 (1991).
- Takai, M. et al. *Biochem. biophys. Res. Commun.* **184**, 953–959 (1992).
- James, A. F. et al. *Cardiovasc. Drug Rev.* **11**, 253–270 (1993).
- Sokolovsky, M. *Neurochemistry* **59**, 809–821 (1992).
- Hilal-Dandan, R., Urasawa, K. & Branton, L. L. *J. biol. Chem.* **267**, 10620–10624 (1992).
- Jones, L. G., Rozich, J. D., Tsutsui, H. & Cooper, G. *Am. J. Physiol.* **263**, H1447–H1454 (1992).
- Taylor, S. J., Chae, H. Z., Rhee, S. G. & Exton, J. H. *Nature* **350**, 516–518 (1991).
- Camps, M. et al. *Nature* **360**, 684–686 (1992).
- Vogelsang, M., Broede-Sitz, A., Schäfer, E., Zerkowski, H.-R. & Brodde, O.-E. *J. cardiovasc. Pharmacol.* **23**, 344–347 (1994).
- Ishikawa, T., Yanagisawa, M., Kimura, S., Goto, K. & Masaki, T. *Am. J. Physiol.* **255**, H970–H973 (1988).
- Wang, J., Fleming, K. & Morgan, J. P. *Am. J. Physiol.* **265**, H2168–H2174 (1993).
- Kohmoto, O. et al. *Am. J. Physiol.* **265**, H793–H800 (1993).
- Lauer, M. R., Gunn, M. D. & Clusin, W. T. *J. Physiol., Lond.* **445**, 729–747 (1992).
- Tohse, N., Hattori, Y., Nakaya, H., Endou, M. & Kanno, M. *Br. J. Pharmacol.* **99**, 437–438 (1990).
- Reid, J. J., Wong-Dusting, H. K. & Rand, M. J. *Eur. J. Pharmacol.* **168**, 93–95 (1989).
- Kim, D. *Circ. Res.* **69**, 250–255 (1991).
- Ono, K. et al. *Nature* **370**, 301–304 (1994).
- Horie, M., Hwang, T.-C. & Gadsby, D. C. *J. Physiol., Lond.* **455**, 235–246 (1992).

ACKNOWLEDGEMENTS. We thank A. Noma, S. Sasayama (Faculty of Medicine, Kyoto Univ.), T. Okada and H. Hofstetter (IRL, Ciba-Geigy Japan) for their encouragement and support. This work was supported by Grants-in-Aid from the Ministry of Science, Education and Culture, Japan. L.-H.X. is the recipient of a scholarship from the Ministry of Science, Education and Culture, Japan.

Endothelin-A receptor mediates cardiac inhibition by regulating calcium and potassium currents

Kageyoshi Ono, Gozoh Tsujimoto*,
 Aiji Sakamoto*, Koji Eto*, Tomoh Masaki†,
 Yukihiro Ozaki & Motoyoshi Satake

Division of Chemical Pharmacology and Phytochemistry,
 National Institute of Health Sciences, 1-18-1 Kami-Yohga,
 Setagaya-ku, Tokyo 158, Japan

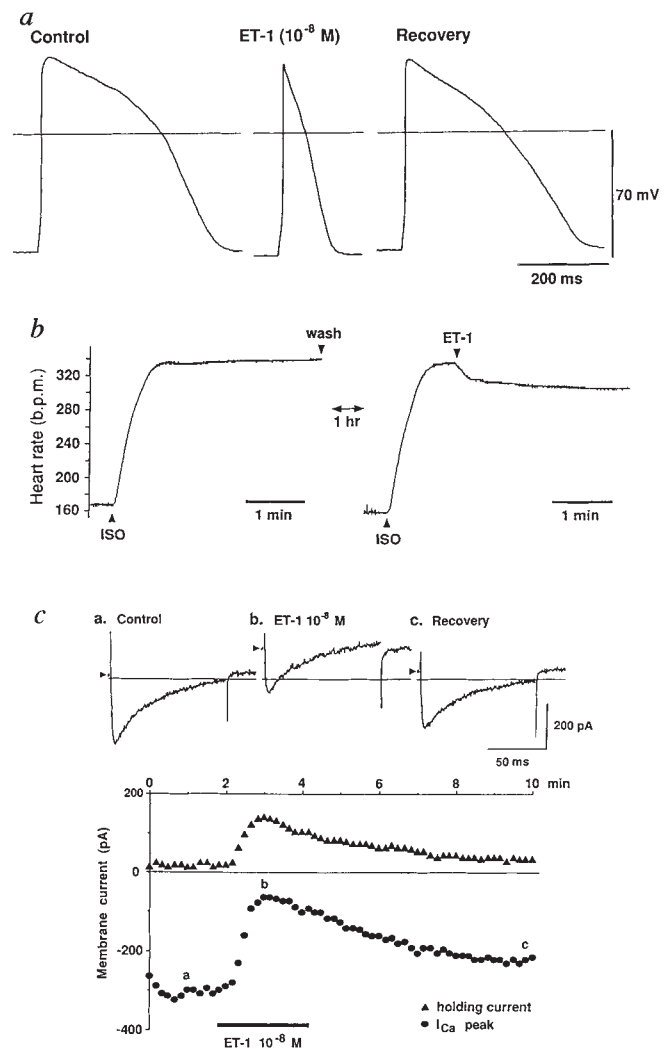
* Division of Molecular and Cellular Pharmacology,
 National Children's Medical Research Centre, 3-35-31 Taishido,
 Setagaya-ku, Tokyo 154, Japan

† Department of Pharmacology, Faculty of Medicine, Kyoto University,
 Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-01, Japan

VOLTAGE-SENSITIVE ion channels play fundamental roles in the regulation of cardiac function by various neurotransmitters^{1,2}. Endothelins³ have strong positive inotropic⁴ and chronotropic⁵ effects, for which recent studies have implicated various intracellular mechanisms^{6,7}. However, very little is known about the underlying ion-channel regulation by the peptide. We report here that endothelin-1 consistently hyperpolarizes the membrane and shortens the duration of the action potential in mammalian atrial myocytes, leading to suppression of electrical excitability of the heart. Endothelin-1, but not endothelin-3, inhibited the L-type calcium current by decreasing cyclic AMP accumulation and activated the muscarinic potassium current by stimulating a pertussis toxin-sensitive GTP-binding protein. Consistent with these results, endothelin-1 strongly reduced the heart rate when it was increased by β -adrenoceptor stimulation. These effects were blocked by an ET_A (endothelin-1-selective) receptor-selective antagonist, BQ123 (refs 8–11). The ET_A receptor-mediated regulation of cardiac ion channels gives new insight into our understanding of the physiological and pathophysiological roles of endothelins in the control of cardiac function.

In single quiescent atrial myocytes isolated from adult guinea-pigs, endothelin-1 (10^{-8} M) hyperpolarized the membrane potential and shortened the duration of the action potential (Fig. 1a). The effect was reversible and the shape of the action potential was restored after continuous washout of the peptide. Correspondingly, endothelin-1 reduced the heart rate when it was increased by a β -adrenergic agonist isoprenaline (ISO, Fig. 1b), which contrasts with the well-known positive chronotropic effect⁵. This suggests that the peptide exerts rather a negative modulatory effect¹² through its electrophysiological action, when the heart is excited by sympathetic stimulation. These findings prompted us to investigate the ionic mechanism of the endothelin-1 actions.

In whole-cell voltage-clamp experiments using both guinea-pig and rabbit atrial myocytes, endothelin-1 at relatively low concentrations ($\sim 3 \times 10^{-9}$ M) both shifted the background hold-

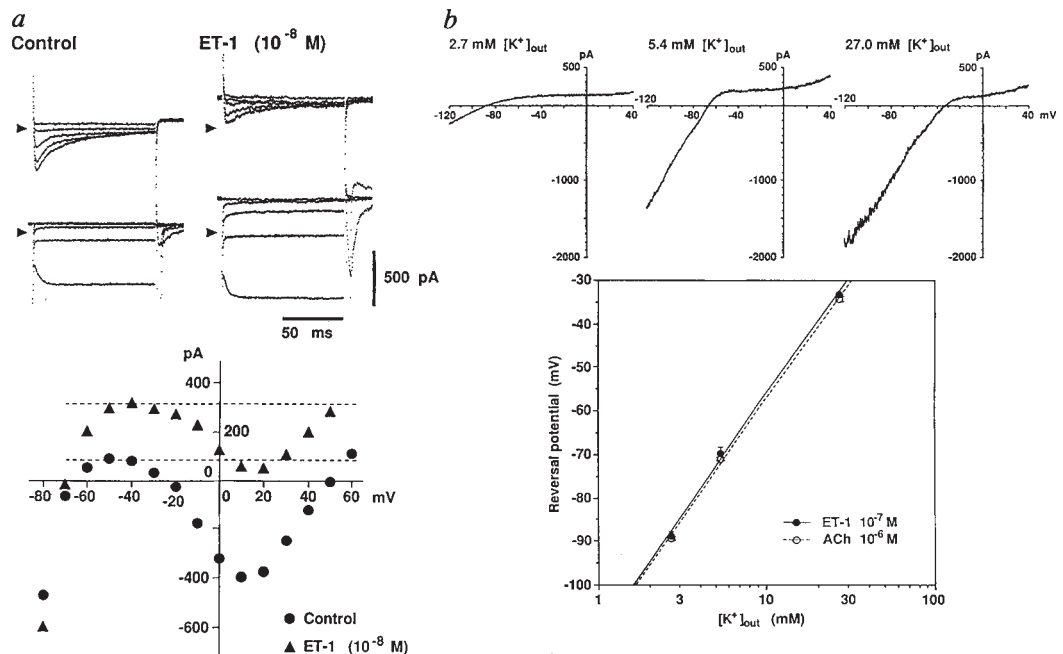


from the current level at the end of test pulse, from 314.4 ± 30.9 pA to 233.4 ± 28.9 pA ($n=5$, $P<0.001$, paired *t*-test), and increased I_{hold} from 22.4 ± 10.8 pA to 132.8 ± 22.9 pA ($n=5$, $P<0.001$, paired *t*-test). Similarly, in rabbit atrial myocytes, ET-1 (10^{-7} M) decreased I_{Ca} to $79.8 \pm 3.4\%$ ($n=3$, $P<0.01$, paired *t*-test) and increased I_{hold} by 70.1 ± 9.6 pA ($n=3$).

METHODS. Male guinea-pigs (350–400 g) or rabbits (2.5–3.0 kg) were injected with heparin (300 U kg^{-1} , i.p.) and deeply anaesthetized with sodium pentobarbital (50 mg kg^{-1} , i.p.), and the whole heart was removed. a, c, Membrane potential and currents were recorded at room temperature from single isolated atrial myocytes with a single electrode patch-clamp technique in the whole-cell configuration²⁶, using a patch-clamp amplifier (EPC-7, List-electronic, Germany). Holding potential = -40 mV. Test potential = $+10$ mV, for 100 ms at 0.1 Hz (guinea-pig) or for 200 ms at 0.2 Hz (rabbit). Data were digitized as described previously^{1,27} at a sampling speed of 4 kHz after filtering at 3 kHz. Myocytes were isolated using collagenase (42 U ml^{-1} ; Yakult, Japan) as described elsewhere^{1,27}. All drugs were dissolved in Tyrode's solution (in mM): NaCl (135.0), KCl (5.4), glucose (5.5), HEPES (5.0), CaCl_2 (1.8), MgCl_2 (0.5), NaH_2PO_4 (0.33), pH 7.4 adjusted with NaOH and bubbled with 100% O_2 , and applied by superfusion at about 2.0 ml min^{-1} in a recording chamber of $\sim 0.5 \text{ ml}$ volume. Electrode solution (in mM): L-aspartic acid mono-K salt (120.0), KCl (30.0), EGTA (5.0), HEPES (5.0), ATP-2Na (4.0), MgCl_2 (1.0), pH 7.2 with KOH. b, Heart rate was counted using MacLab (Analog Digital Instruments, NSW, Australia) by monitoring the spontaneous beatings of the right atrium, with a force-displacement transducer, in an organ bath containing 20 ml modified Krebs solution (in mM): NaCl (113.0), KCl (4.8), glucose (5.5), NaHCO_3 (25.0), CaCl_2 (2.2), MgSO_4 (1.2), KH_2PO_4 (1.2), bubbled with 95% O_2 -5% CO_2 at 37°C . After equilibration for 1 h, 5-min applications of ISO were repeated 3 times with 65-min intervals, and ET-1 was added during the third application of ISO.

FIG. 1 Effects of endothelin-1 (ET-1); on action potential (a), heart rate (b) and related membrane currents (c) in guinea-pig atrium. a, ET-1 (10^{-8} M) hyperpolarized the resting membrane potential from -63.5 ± 1.0 mV to -68.1 ± 0.3 mV ($n=3$, $P<0.05$; paired *t*-test) and decreased APD_{50} (a 50% repolarizing period of the action potential) to $22.2 \pm 0.2\%$ of control ($n=3$, $P<0.001$ paired *t*-test). 'Recovery' 10 min after washout of ET-1. b, Heart rate changes by isoprenaline (ISO; left) or by ISO plus ET-1 (right). ET-1 (10^{-7} M) decreased heart rate to $83.6 \pm 1.9\%$ ($n=6$, $P<0.01$) of the ISO (3×10^{-7} M)-induced maximal elevation. c, Effects of ET-1 (10^{-8} M) on I_{Ca} and I_{hold} . Bottom, time courses of I_{hold} ($\blacktriangle$) and peak I_{Ca} ($\bullet$). Top, I_{Ca} recordings obtained as marked in the bottom panel. ET-1 reduced peak I_{Ca} size, measured

FIG. 2 Effects of endothelin-1 (ET-1) on I_{Ca} and time-independent background current in guinea-pig atrial myocytes. *a*, Top, Actual recordings of I_{Ca} and the background current, obtained at -30 mV to $+10$ mV and -50 mV to -80 mV, respectively, before (left) and after (right) application of ET-1 (10^{-8} M). Arrow heads indicate zero-current level. Bottom, I - V relation. Peak I_{Ca} values recorded at -30 mV to $+60$ mV and the current sizes measured at the end of test pulses of -80 mV to -40 mV are plotted. Dotted lines show $I_{(hold)}$ at -40 mV. Note that ET-1 did not affect the voltage-dependence of activation of I_{Ca} . *b*, Effect of ET-1 (10^{-7} M) on a time-independent component of the background current. Top, Instantaneous I - V relation of ET-1-induced current measured at three different extracellular K^+ concentrations ($[K^+]_{out}$). Bottom, Correlation between $[K^+]_{out}$ and the reversal potential (E_{rev}) of the ET-1-induced current (●). Results obtained in the same cells with ACh (10^{-6} M) applied after washout of ET-1 (○) are also shown. Each symbol represents the mean of 4 to 6 experiments with s.e.m. Straight lines were obtained by the least-squares linear regression, giving a slope of 55.0 mV per (log unit) for ET-1 ($r^2=0.998$) and 54.7 mV per (log unit) for ACh ($r^2=0.999$). Lower concentrations of ET-



1, such as 10^{-8} M, gave qualitatively identical results, both in the current-voltage profiles and in the $[K^+]_{out}$ -dependence of E_{rev} . **METHODS.** *a*, Square pulses (100 ms) were applied from -80 mV to $+60$ mV with 10-mV increments at 1/5 Hz. Holding potential = -40 mV. *b*, Instantaneous I - V relationships obtained by applying voltage ramp, from $+50$ mV to -120 mV in 5 s, were compared before and after application of ET-1, and the difference current between these two are shown as the ET-1-induced current.

ing current $I_{(hold)}$ to the outward direction and decreased the high-threshold calcium current, I_{Ca} (Fig. 1c). The current-voltage (I - V) relationship for I_{Ca} in Fig. 2a illustrates the effect of endothelin-1 (10^{-8} M). Endothelin-1 decreased I_{Ca} without noticeably affecting its voltage-dependence or the time courses of either activation or inactivation, and increased the time-independent background current, as measured at membrane potentials negative to -50 mV and as indicated by the outward shift of $I_{(hold)}$.

The endothelin-1-induced fraction of the background current showed inward rectification and was highly selective to K^+ ions (Fig. 2b). It was abolished by 0.1 mM $BaCl_2$ but was unaffected by ryanodine (10^{-6} M), which blocks Ca^{2+} release from the sarcoplasmic reticulum, or atropine (10^{-6} M), although the effect of endothelin-1 resembled that of acetylcholine (ACh)¹⁴. Endothelin-1 did not affect the background current in ventricular myocytes even at very high concentration (3×10^{-7} M, data not shown). These results indicate that endothelin-1 activates a population of K^+ currents quite similar to the muscarinic potas-

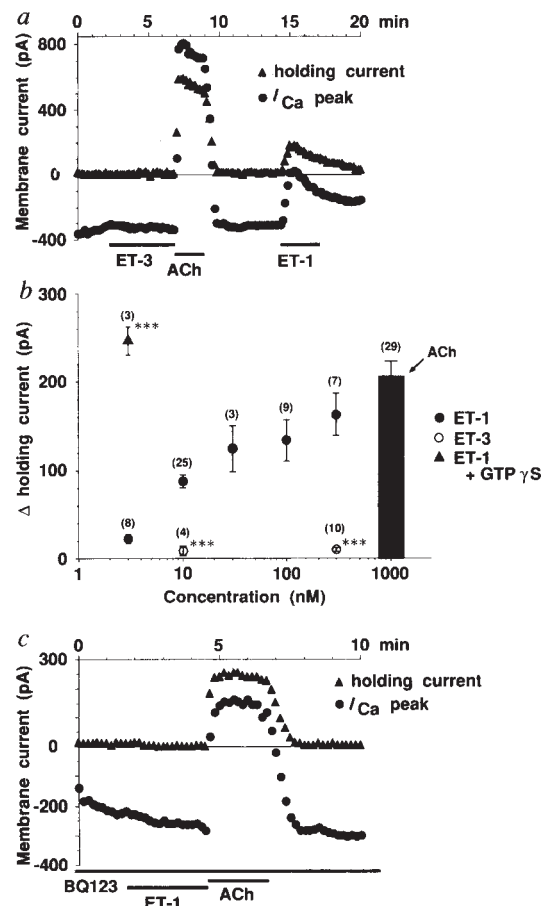
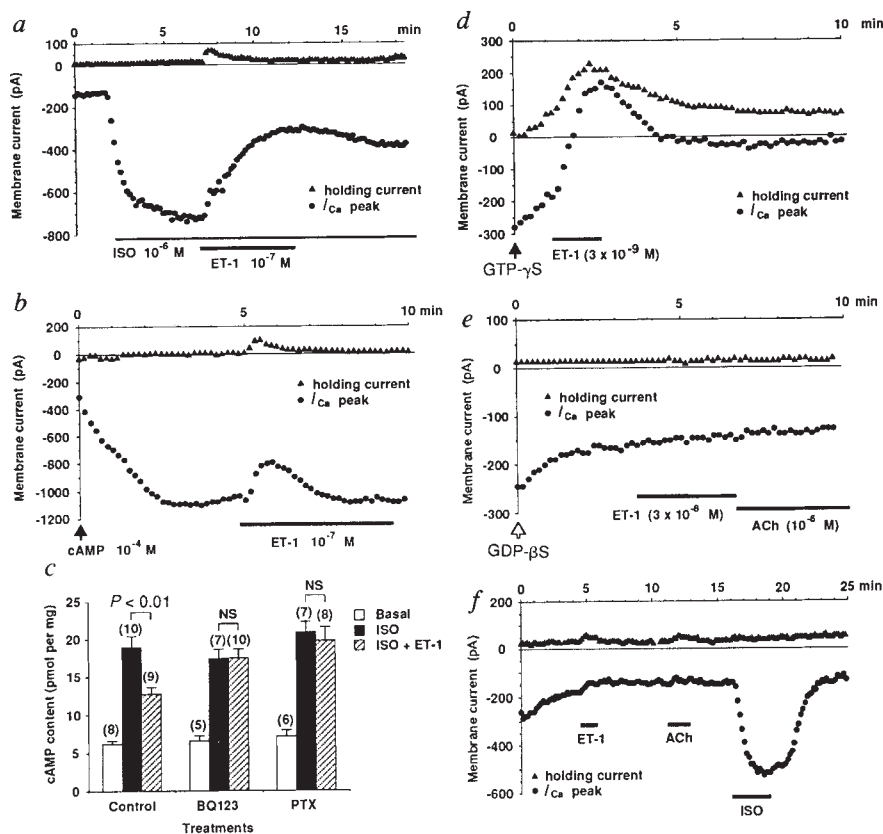


FIG. 3 Studies on endothelin-receptor-subtype responsible for the effect of ET-1. *a*, Comparison of the effects of ET-1 (10^{-8} M), ET-3 (10^{-8} M) and ACh (10^{-6} M). Experimental condition was the same as in Fig. 1c. Drugs were applied as indicated by bars. *b*, Concentration-response relation of ET-1 (●) and ET-3 (○) for the increase in $I_{(hold)}$ at -40 mV. A single concentration of ETs was tested for each cell. Effect of ET-1 (▲) (3×10^{-9} M) recorded by intracellularly applying GTP- γ S²⁷ (5×10^{-5} M, see Fig. 4d). Each point represents mean $\pm$ s.e.m. in parentheses are the number of experiments. The bar graph shows the maximal effect of ACh (10^{-6} M). ***, Significantly different ($P < 0.001$, Student's *t*-test) from ET-1 (10 and 300 nM) or from GTP- γ S-untreated cells (3 nM). *c*, Blockade of the effects of ET-1 (10^{-8} M) by BQ123 (10^{-6} M) with which myocytes were preincubated for 5 min. Note that the effect of ACh (10^{-6} M) was unaffected; it greatly increased $I_{K(ACh)}$ and significantly inhibited I_{Ca} .

Fig. 4 a–c, Involvement of cAMP in the ET-1-induced inhibition of I_{Ca} ; $I_{(hold)}$ ($\blacktriangle$), peak I_{Ca} ($\bullet$). a, b, Atropine (10^{-6} M) was present throughout the experiments. a, Inhibition by ET-1 (10^{-7} M) of ISO (10^{-6} M)-stimulated I_{Ca} . Note that I_{Ca} partly recovered after washout of ET-1. b, Effect of ET-1 (10^{-7} M) on I_{Ca} in myocytes internally perfused with cAMP (10^{-4} M). Note that on recovery of $I_{(hold)}$, I_{Ca} was fully restored. Substituting Cs^{+} for both extra- and intracellular K^{+} totally abolished the effect of ET-1 in this condition. c, Inhibition by ET-1 (10^{-7} M) of ISO (3×10^{-7} M)-stimulated cAMP accumulation in atrial tissues. Left, The cAMP contents were 6.2 ± 0.5 pmol mg^{-1} protein (basal, $n=8$), 19.0 ± 1.5 pmol mg^{-1} protein (ISO, $n=10$, $P<0.01$) and 12.7 ± 1.0 (ISO plus ET-1, $n=9$, $P<0.01$, compared with ISO alone). Isobutylmethylxanthine (10^{-4} M) did not affect the ET-1 action (data not shown), indicating that ET-1 acts by inhibiting cAMP production. Middle, BQ123 (10^{-6} M, pretreated for 20 min) blocked the effect of ET-1, whereas it did not affect either (1) basal cAMP and heart rate (control, 6.2 ± 0.5 pmol mg^{-1} protein, $n=8$, and 150 ± 4 beats per min (b.p.m.), $n=11$; BQ123, 6.6 ± 0.7 pmol mg^{-1} protein, $n=5$, and 146 ± 8 b.p.m., $n=4$) or (2) ISO-induced elevation of cAMP and heart rate (control, 19.0 ± 1.5 pmol mg^{-1} protein, $n=10$, and 307 ± 6 b.p.m., $n=11$; BQ123, 17.4 ± 1.3 pmol mg^{-1} protein, $n=7$, and 302 ± 7 b.p.m., $n=4$). Right, Abolition of the effect of ET-1 in PTX-treated tissues. Each bar indicates mean $\pm$ s.e.m. In parentheses are the number of experiments. NS, Not significantly different. d, e, Effects of intracellularly perfused GTP analogues on the ET-1-induced responses. d, GTP- γ S (5×10^{-5} M) by itself gradually increased $I_{(hold)}$ ($\blacktriangle$) and decreased I_{Ca} ($\bullet$) immediately after the whole-cell recording started. ET-1 (3×10^{-9} M) greatly accelerated the change in $I_{(hold)}$ and inhibited I_{Ca} dramatically, in an almost irreversible manner. e, GDP- β S (2×10^{-3} M) blocked the effects of both ET-1 (10^{-8} M) and ACh (10^{-6} M). f, Attenuation of the effects of ET-1 (10^{-8} M) and ACh (10^{-6} M) in myocytes prepared from PTX-treated (5 μ g per 100 g body weight, i.v., 4 days) guinea-pigs, where the effect of ISO (10^{-6} M), mediated



ated by a stimulatory GTP-binding protein G_s , remained unaffected. METHODS. cAMP content: after incubating in Tyrode's solution for 1 h at 37 $^{\circ}$ C, atrial tissue was incubated with ISO (3×10^{-7} M) for 4.5 min and frozen immediately in liquid nitrogen. For another batch of tissues, ET-1 (10^{-7} M) was added 30 s after the application of ISO, when heart rate usually reaches maximal elevation (see Fig. 1b), and incubated for further 4 min, by which time the ET-1-induced reduction in heart rate reaches a steady level. The cAMP content was measured by radioimmunoassay and the protein content by the Lowry method.

sium current, $I_{K(ACh)}$ (refs 14, 15), in adult mammalian atrial cells under physiological conditions. A similar activation of $I_{K(ACh)}$ by endothelin-3 has recently been demonstrated in cultured neonatal rat heart cells¹⁶. These electrophysiological actions of endothelin-1 could account for the peptide-induced changes in the shape of the action potential and therefore may play an important role in the regulation of the electrical excitability of cardiac muscle.

We further examined the physiological significance of the electrophysiological actions of endothelin-1 by characterizing their concentration-dependency. Endothelin-1, a non-selective agonist^{9-11,17,18}, but not endothelin-3, an ET_B -receptor-selective agonist^{10,11}, strongly increased $I_{K(ACh)}$, as indicated by the outward shift of $I_{(hold)}$ (Fig. 3a). The concentration-response curve in Fig. 3b shows that the maximal effect was obtained at about 3×10^{-7} M, which was almost comparable to that of acetylcholine (10^{-6} M). The effective concentration for half-maximum response (EC_{50}) was about 2×10^{-8} M. In contrast, endothelin-3 was not effective. Furthermore, an ET_A -receptor-selective antagonist, BQ123 (ref. 8; 10^{-6} M), completely blocked the effects of endothelin-1 (Fig. 3c). In contrast, an ET_B -receptor-selective antagonist, RES-701-1 (ref. 19; 5×10^{-5} M), was without effect (data not shown). Recently, both ET_A and ET_B receptors have been shown to coexist in the rat heart, using binding assay²⁰ and *in situ* hybridization²¹. Using [¹²⁵I]ET-1 and BQ123, as radioligand and competitor respectively, we have found that adult guinea-pig atrium expresses almost equal populations of both ET_A and ET_B receptors ($ET_A:ET_B=51:49$,

unpublished observation). The endothelin-1-induced electrophysiological responses are thus most probably mediated by the ET_A receptor.

Further experiments attempted to obtain information about the intracellular mechanism(s) of the endothelin-1-induced electrophysiological responses, with particular focus on the involvement of cAMP and GTP-binding proteins. Endothelin-1 (10^{-7} M) strongly decreased I_{Ca} when it was maximally increased by isoprenaline (10^{-6} M, Fig. 4a). In contrast, endothelin-1 failed to decrease I_{Ca} when it was increased by perfusing cAMP (10^{-4} M) intracellularly (Fig. 4b), although it transiently shifted the current trace to the outward direction because of the activation of $I_{K(ACh)}$ (ref. 14). Moreover, endothelin-1 noticeably decreased tissue cAMP content which was elevated by isoprenaline (3×10^{-7} M), by inhibiting cAMP production (Fig. 4c), as has been reported recently in the rat ventricle and human atrium^{22,23}. These results indicate that endothelin-1 decreases I_{Ca} by inhibiting cAMP production.

Intracellularly applied GTP- γ S (5×10^{-5} M) strongly potentiated the effects of endothelin-1 on both $I_{K(ACh)}$ and I_{Ca} (Fig. 4d). In contrast, GDP- β S (2×10^{-3} M) blocked the effect of the peptide, where acetylcholine (10^{-6} M) also failed to stimulate $I_{K(ACh)}$ or to decrease I_{Ca} (Fig. 4e). This GTP-analogue-dependence of endothelin-1 action also confirms that the endothelin-1-induced background current is $I_{K(ACh)}$ (ref. 15). Moreover, pretreatment of guinea-pigs with pertussis toxin (PTX) noticeably attenuated the effects of endothelin-1 (Fig. 4f); in such cells, membrane protein(s) of relative molecular

mass 40,000–41,000 (40–41K) had been ADP-ribosylated, so that *in vitro* incorporation by PTX of [³²P]ADP-ribose into the protein was attenuated by 79 ± 9% (*n* = 5; *P* < 0.01). These results strongly indicate that endothelin-1 induces the electrophysiological responses through a PTX-sensitive GTP-binding protein, probably G_i or G_o.

Endothelins are released in various stressful conditions such as cardiac infarction and congestive heart failure^{6,24}, where sympathetic tone is extensively elevated. Therefore, interaction of endothelins with sympathetic neurotransmitters²⁵ may play a very important role in both physiological and pathophysiological states. Using β-adrenoceptor stimulation as an example of such conditions, we found that endothelin-1, through the ET_A receptors, exerts a negative modulatory action by interacting with certain ion channels. Our results provide the first direct evidence that the ET_A receptors mediate hyperpolarization and shortening of the action potential duration by inhibiting *I*_{Ca} and stimulating *I*_{K(ACh)} in adult mammalian heart under physiological conditions. The effects are coupled to a PTX-sensitive GTP-binding protein and the inhibition of cAMP accumulation. These ET_A-receptor-mediated electrophysiological actions of endothelin-1 may give important insights into our understanding of the physiological and pathophysiological regulation of heart functions. □

Received 28 March; accepted 22 June 1994.

1. Ono, K., Delay, M., Nakajima, T., Irisawa, H. & Giles, W. *Nature* **340**, 721–724 (1989).
2. Reuter, H. A. *Rev. Physiol.* **41**, 413–424 (1979).
3. Yanagisawa, M. et al. *Nature* **332**, 411–415 (1988).

4. Ishikawa, T., Yanagisawa, M., Kimura, S., Goto, K. & Masaki, T. *Am. J. Physiol.* **255**, H970–H973 (1988).
5. Ishikawa, T., Yanagisawa, M., Kimura, S., Goto, K. & Masaki, T. *Pflügers Arch. ges. Physiol.* **413**, 108–110 (1988).
6. Masaki, T., Yanagisawa, M. & Goto, K. *Med. Res. Rev.* **12**, 391–421 (1992).
7. Aramori, I. & Nakanishi, S. *J. Biol. Chem.* **267**, 12468–12474 (1992).
8. Ihara, M. et al. *Life Sci.* **50**, 247–255 (1992).
9. Arai, H., Hori, S., Aramori, I., Ohkubo, H. & Nakanishi, S. *Nature* **348**, 730–732 (1990).
10. Sakamoto, A. et al. *J. Biol. Chem.* **268**, 8547–8553 (1993).
11. Sakurai, T. et al. *Nature* **348**, 732–735 (1990).
12. Franco-Cereceda, A., Matran, R., Lou, Y.-P. & Lundberg, J. M. *Acta physiol. scand.* **138**, 539–547 (1990).
13. Tohse, N., Hattori, Y., Nakaya, H., Endou, M. & Kanno, M. *Br. J. Pharmacol.* **99**, 437–438 (1990).
14. Hartzell, H. C. & Simmons, M. A. *J. Physiol., Lond.* **389**, 411–422 (1987).
15. Pfaffinger, P. J., Martin, J. M., Hunter, D. D., Nathanson, N. M. & Hille, B. *Nature* **317**, 536–538 (1985).
16. Kim, D. *Circ. Res.* **69**, 250–255 (1991).
17. Sakamoto, A. et al. *Biochem. biophys. Res. Commun.* **178**, 656–663 (1991).
18. Inoue, A. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2863–2867 (1989).
19. Tanaka, T. et al. *Molec. Pharmacol.* **45**, 724–730 (1994).
20. Ishikawa, T. et al. *Circ. Res.* **69**, 918–926 (1991).
21. Hori, S., Komatsu, Y., Shigemoto, R., Mizuno, N. & Nakanishi, S. *Endocrinology* **130**, 1885–1895 (1992).
22. Hilal-Dandan, R., Urasawa, K. & Brunton, L. L. *J. Biol. Chem.* **267**, 10620–10624 (1992).
23. Vogelsang, M., Broede-Sitz, A., Schäfer, E., Zerkowski, H.-R. & Brodde, O.-E. *J. cardiovasc. Pharmacol.* **23**, 344–347 (1994).
24. Wennmalm, A., Benthin, G., Karwatowska-Prokopczuk, E., Lundberg, J. & Petersson, A.-S. *J. Physiol., Lond.* **435**, 163–173 (1991).
25. Reid, J. J., Lieu, A. T. & Rand, M. J. *Eur. J. Pharmacol.* **194**, 173–181 (1991).
26. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
27. Ono, K. & Giles, W. R. *J. Physiol., Lond.* **436**, 195–217 (1991).

ACKNOWLEDGEMENTS. We thank W. R. Giles (Departments of Medical Physiology and Medicine, University of Calgary, Canada) for his helpful discussion and M. Ihara and M. Yano of Banyu Pharmaceutical Co. Ltd, Japan and Y. Matsuda of Kyowa Hakkō Kogyo Co., Japan for kindly providing BQ123 and RES-701-1, respectively. This work was supported in part by the Alberta Heritage Foundation for Medical Research, Canada, and Research Grant for Encouragement of Young Researchers, Ministry of Science and Education, Japan. K.O. thanks Nissan Science Foundation, Japan for financial support.

Mice with targeted disruptions in the paralogous genes *hoxa-3* and *hoxd-3* reveal synergistic interactions

Brian G. Condie & Mario R. Capecchi*

Howard Hughes Medical Institute, Department of Human Genetics, University of Utah School of Medicine, Salt Lake City, Utah 84112, USA

THE *Hox* genes encode transcription factors which mediate the formation of the mammalian body plan along the anteroposterior and appendicular axes^{1–15}. Paralogous *Hox* genes within the separate linkage groups are closely related with respect to DNA sequence and expression^{16,17}, suggesting that they could have at least partially redundant functions. We showed previously that mice homozygous for independent targeted disruptions in the paralogous genes *hoxa-3* and *hoxd-3* had no defects in common^{1,8}. But our current analysis of double mutants has revealed strong, dosage-dependent interactions between these genes. We report here that in *hoxd-3*[−] homozygotes the first cervical vertebra, the atlas, is homeotically transformed to the adjacent anterior structure. Unexpectedly, in double mutants, rather than observing a more extensive homeotic transformation, the entire atlas is deleted. These observations are interpreted in terms of a model in which these *Hox* genes differentially regulate the proliferation rates of the appropriate sets of precursor cells.

The absence of overlapping abnormalities in *hoxa-3*[−] and *hoxd-3*[−] mutant mice does not imply that these two genes do not interact. To investigate such interactions, we generated mice that are simultaneously mutant for both genes. Because *hoxa-3*[−] and *hoxd-3*[−] homozygotes usually die shortly after birth (*hoxa-3*[−]/*hoxa-3*[−] mice from cardiovascular dysfunction and

hoxd-3[−]/*hoxd-3*[−] mice from apparent accidental cervical dislocation as a result of a destabilized craniocervical joint), double homozygotes were generated by intercrossing compound heterozygotes. From these intercrosses, 284 full-term progeny were produced. Of these, 14 were double homozygotes, very close to the expected frequency of 1/16, indicating that such mice do not die during embryogenesis.

Figure 1 shows lateral views of skeleton preparations of wild-type mice, *hoxa-3*[−] homozygotes, *hoxd-3*[−] homozygotes and double mutants. The craniocervical joint of *hoxa-3*[−] homozygotes is indistinguishable from that of wild-type animals (compare Fig. 1a and b). The *hoxd-3* mutation alone results in a remodelling of the bones in the craniocervical joint. In these mice, the anterior arch of the atlas forms an extension of the basioccipital bone, the atlas lateral masses are remodelled to resemble and are often fused to the exoccipitals, the axis is remodelled to resemble the atlas, the dens and the superior facets of the axis are deleted and the lateral foramina normally associated with the atlas are now attached to the axis (Fig. 1c and ref. 8). The double mutant mouse, however, shows a more severe phenotype. The entire atlas has been deleted, and only a small piece of cartilage remains in its place (Fig. 1d). This deletion of

FIG. 2 The deletion of the atlas depends on the dosage of *Hoxa-3* mutant alleles in the *hoxd-3*[−] homozygote. Lateral views of cleared skeleton preparations of newborn mice. a, The craniocervical region in a wild-type newborn (a3+/+; d3+/+). The atlas (at) and axis (ax) are indicated. The arrow indicates the ossification centre in the anterior arch of the atlas. b, Same view of a *hoxd-3* heterozygote (a3+/+; d3+/−). c, The atlas is further altered in this compound heterozygote (a3+/−; d3+/−). The ossification centre in the anterior arch of the atlas has failed to form (arrow). d, A *hoxd-3* homozygote (a3+/+; d3−/−) with a nearly complete homeotic transformation of the atlas. This is an example of the most extreme *hoxd-3* phenotype. e, Compared with the *hoxd-3*[−] homozygote, the atlas is more severely affected in the (*hoxa-3*[−]/*hoxa-3*[−]; *hoxd-3*[−]/*hoxd-3*[−]) mouse. f, A double-mutant newborn (a3−/−; d3−/−). The atlas is deleted and cartilaginous fusions between the neural arches of the axis and other cervical vertebrae are indicated (arrow). Scale bar, 1.0 mm.

* To whom correspondence should be addressed.

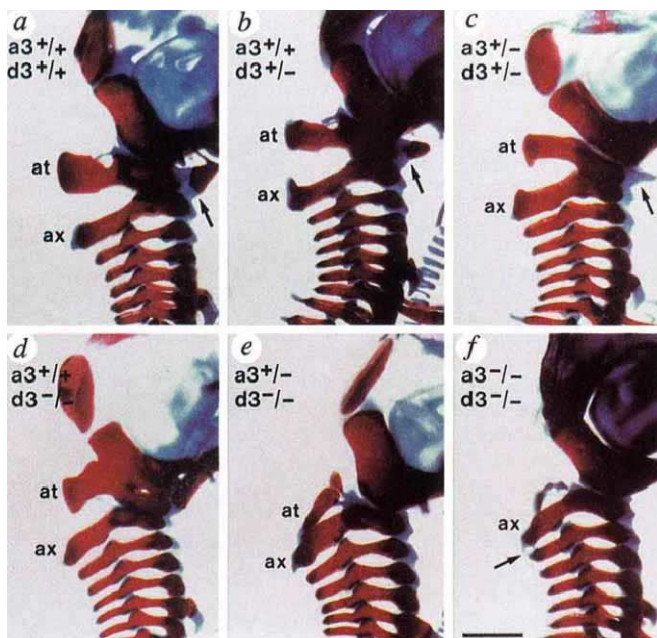


FIG. 1 Additional skeletal defects in (*hoxa-3⁻/hoxa-3⁻; hoxd-3⁻/hoxd-3⁻*) mice. Lateral views of cleared skeleton preparations of newborn mice. **a**, A wild-type (*a3^{+/+}; d3^{+/+}*) newborn, the first cervical vertebra or atlas (**at**) and the second cervical vertebra or axis (**ax**) are labelled. **b**, A *Hoxa-3* homozygote (*a3^{-/-}; d3^{+/+}*). **c**, A *Hoxd-3* homozygote (*a3^{+/+}; d3^{-/-}*) displaying the characteristic homeotic transformation of the atlas (**at**). **d**, A *hoxa-3*, *hoxd-3* double homozygote (*a3^{-/-}; d3^{-/-}*) showing the complete absence of the atlas and remodelling of the axis (**ax**). Scale bar, 1 mm.

METHODS. Newborn skeleton preparations were made as described²¹.

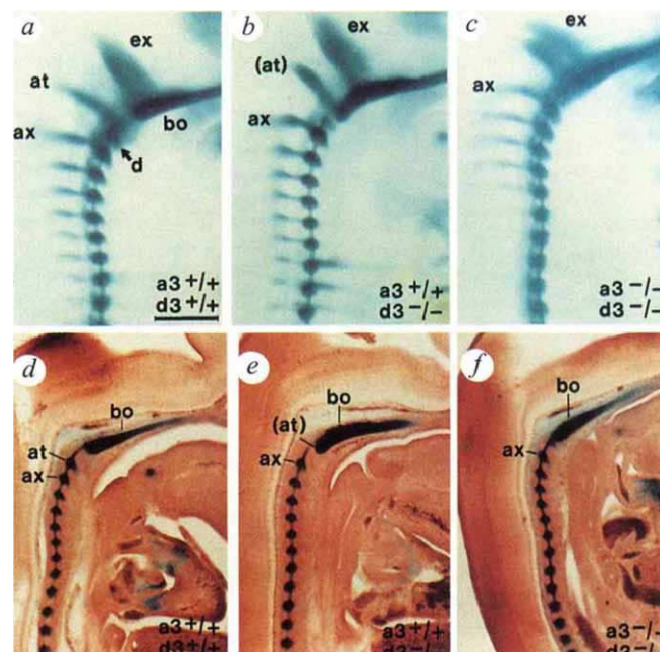
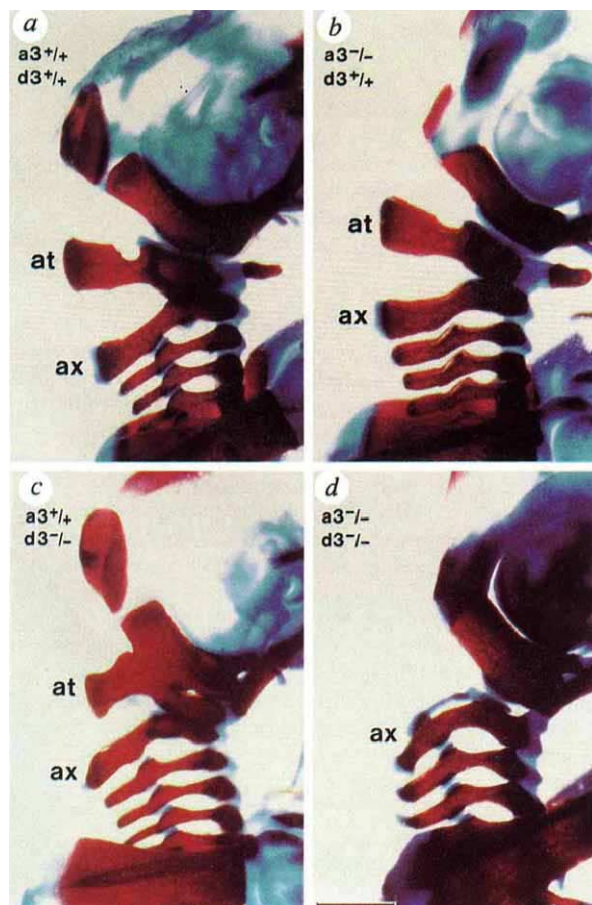


FIG. 3 The pattern of the cartilage condensations in E13 embryos. **a-c**, Lateral views of cleared preparations of mouse embryos. **a**, A wild-type embryo (*a3^{+/+}; d3^{+/+}*). The cartilage anlagen of the basioccipital (**bo**), exoccipital (**ex**), atlas (**at**), axis (**ax**) and dens (**d**) are indicated. **b**, A *hoxd-3* homozygous embryo (*a3^{+/+}; d3^{-/-}*). **c**, A double-mutant embryo (*a3^{-/-}; d3^{-/-}*). The atlas is deleted and the cartilage anlage of the axis is indicated. **d-f**, Midsagittal thick sections (100 μ m) of the embryos pictured in **a-c**. **d**, The wild-type embryo (*a3^{+/+}; d3^{+/+}*), the cartilage anlagen for the basioccipital (**bo**), atlas (**at**) and axis (**ax**) are indicated. **e**, The *hoxd-3* homozygote (*a3^{+/+}; d3^{-/-}*). **f**, The double-mutant embryo (*a3^{-/-}; d3^{-/-}*). The assignment of the cervical vertebrae anlagen in wild-type and mutant mice was confirmed by counting rostrally from the first thoracic vertebra anlage. Scale bar, 0.5 mm.

METHODS. Embryos were collected on embryonic day 13 (morning of the plug is 0.5 days), stained with Alcian blue and cleared as described²². Stained embryos were removed from benzyl benzoate/benzyl alcohol into methanol, rehydrated in PBS, embedded in 3% agarose and sectioned at 100 μ m on a vibratome. Sections were dried onto gelatin subbed slides, counterstained with nuclear Fast red and mounted in 90% glycerol/PBS.

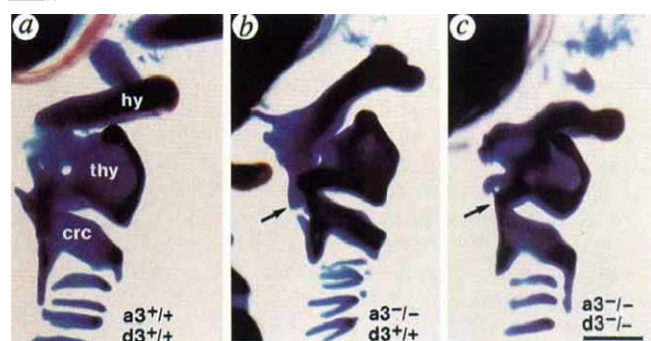


FIG. 4 The *Hoxa-3⁻* defects are more severe in (*hoxa-3⁻/hoxa-3⁻; hoxd-3⁻/hoxd-3⁻*) mice. Lateral views of skeleton preparations of newborn mice. **a**, The laryngeal cartilages in a wild-type (*a3^{+/+}; d3^{+/+}*) newborn mouse. The hyoid bone (**hy**), thyroid (**thy**) and cricoid (**crc**) cartilages are indicated. **b**, The laryngeal cartilages of a *hoxa-3* (*a3^{-/-}; d3^{+/+}*) mutant. The arrow indicates the inferior horn of the thyroid cartilage. **c**, Lateral view of a double homozygote (*a3^{-/-}; d3^{-/-}*). Remodelling has changed the appearance of the hyoid and thyroid cartilages. The dorsal side of the thyroid cartilage (left side) is altered because of the deletion of cartilage and the inferior horn of the thyroid cartilage is missing (arrow). Scale bar, 0.5 mm.

the atlas was observed in all six of the double mutant skeletons examined. The axis has also been remodelled to appear more like a 'generic' cervical vertebra.

We reported previously⁸ that at low penetrance, mice heterozygous for the *hoxd-3*⁻ mutation have mild abnormalities in the craniocervical joint (Fig. 2b). The anterior arch of the atlas is reduced in size and has small, unilateral fusions with the basioccipital bone. These results suggested that the dosage of the *hoxd-3* protein is an important factor in determining the severity of the abnormalities, and raised the possibility that the interactions between *hoxa-3* and *hoxd-3* might also be quantitative (that is, dosage-dependent). Figure 2 shows that the severity of the craniocervical joint abnormality indeed increases as the number of wild-type copies of *hoxa-3* and *hoxd-3* decreases. Compound heterozygotes (*hoxa-3*⁻/*hoxa-3*⁺; *hoxd-3*⁻/*hoxd-3*⁺) show a more severe mutant phenotype at higher penetrance (45% versus 15%) than mice heterozygous for the *hoxd-3*⁻ mutation alone (compare Fig. 2b and c). The anterior arch of the atlas is more reduced and the atlas lateral masses are more extensively remodelled than in *hoxd-3*⁻/*hoxd-3*⁺ mice; also note the absence of the lateral foramina. In addition, whereas the *hoxd-3*⁻ heterozygotes always have a dens, in the most strongly affected compound heterozygotes the dens is deleted. Furthermore, (*hoxa-3*⁻/*hoxa-3*⁺; *hoxd-3*⁻/*hoxd-3*⁺) mice have a phenotype intermediate between that of the (*hoxa-3*⁺/*hoxa-3*⁺; *hoxd-3*⁻/*hoxd-3*⁻) and (*hoxa-3*⁻/*hoxa-3*⁻; *hoxd-3*⁻/*hoxd-3*⁻) mice (Fig. 2d-f). Parts of the atlas are progressively deleted as one and then two wild-type copies of *hoxa-3* are removed from the *hoxd-3*⁻ homozygous mutant background.

The bones of the vertebral column are formed by the endochondral pathway in which a cartilaginous prototype is laid down before bone formation. The effect of the *hoxd-3*⁻ and *hoxa-3*⁻ mutations can be seen as the precartilaginous condensations are formed. Figure 3a shows a lateral view of an E13 wild-type embryo stained to reveal the cartilage anlagen. Figure 3d shows a sagittal section of the same embryo counterstained with nuclear fast red. The precartilaginous condensations that give rise to the basioccipital, the exoccipitals, the atlas, the axis and the dens are evident. In the *hoxd-3*⁻ homozygotes (Fig. 3b, e), the cartilage anlage that would form the atlas is shifted towards the base of the skull and fused along its ventral aspects with the basioccipital anlage (Fig. 3e). The shape of the axis anlage has changed to resemble more closely the atlas anlage and the dens anlage is severely reduced or absent. In the double mutant (Fig. 3c, f), only a vestige of the atlas anlage remains ventral to the notocord. Thus, the change of the *hoxd-3* phenotype in the double mutants is evident at this embryonic stage.

Figure 4 shows that the *hoxa-3* mutant phenotype is similarly exacerbated in mice also mutant for *hoxd-3*. In *hoxa-3* mutants, the lesser horns of the hyoid bone are reduced or missing and the greater horn of the hyoid is fused with the superior horn of the thyroid cartilage (Fig. 4b). In *hoxd-3*⁻ homozygotes these elements are not affected⁸. In the double mutant, the hyoid is reduced in size and fuses closer to the body of the thyroid cartilage (Fig. 4c). In addition to these changes, cartilage is missing dorsally, creating gaps and holes in the lamina of the thyroid cartilage. Also, the inferior horns of the thyroid cartilage are deleted. These deletions reduce the size of the thyroid cartilage and give it a very different appearance from those of the wild-type and *hoxa-3*⁻ homozygous mutants (Fig. 4a-c). The exacerbation of the *hoxa-3* phenotype by the *hoxd-3*⁻ mutation is also dosage-dependent. For example, although neither *hoxa-3* and *hoxd-3* heterozygotes nor *hoxd-3* homozygotes show defects in the hyoid/thyroid cartilages, compound heterozygotes (*hoxa-3*⁻/*hoxa-3*⁺; *hoxd-3*⁻/*hoxd-3*⁺) exhibit fusions between the hyoid and thyroid cartilages (data not shown).

Not all of the *hoxa-3*⁻ phenotypes are exacerbated by the *hoxd-3*⁻ mutation. For example, the defects of the heart and its great vessels are the same in *hoxa-3*⁻ homozygotes and in the double mutants. Also, the extent of truncation of the soft palate

and the reduction in the thyroid and submaxillary glands in *hoxa-3*⁻/*hoxa-3*⁻ and (*hoxa-3*⁻/*hoxa-3*⁻; *hoxd-3*⁻/*hoxd-3*⁻) mice appear to be the same. In contrast some *hoxa-3*⁻ homozygotes have remnants of a thymus, whereas all double mutants are athymic. Also, the phenotype of the epiglottis in double mutants is more severe than in *hoxa-3*⁻ homozygotes.

Analysis of mice with mutations in both *hoxa-3* and *hoxd-3* has revealed strong genetic interactions between these two genes. The double-mutant mice have more severe abnormalities than predicted from the sum of the individual mutant phenotypes. In the double mutants the entire atlas was deleted rather than being more extensively transformed into the adjacent anterior bones. The extent of deletion of the atlas in *hoxd-3*⁻/*hoxd-3*⁻ mice increased progressively with the loss of one and then two copies of the wild-type *hoxa-3* gene. The deletion of the atlas precursors could be detected in E13 embryos, indicating that the defects seen in the newborn animals reflect changes in the early patterning of the cartilage anlagen for these bones.

To explain the observation that in *hoxd-3*⁻/*hoxd-3*⁻ mice parts of the atlas and axis were homeotically transformed to resemble more anterior structures, whereas other parts were deleted, we suggested a model whereby the *hoxd-3* product differentially regulates the rate of proliferation of precursor cells⁸. Our finding that the entire atlas is deleted in the double mutant lends support to this model. In mice carrying both mutations, we propose that all of the precursor cells responsible for the formation of the atlas fail to proliferate, and thus cannot form the atlas.

In the context of this model, note that *hoxb-8* is transcriptionally activated in mouse myeloid leukaemia WEHI-3B cells, and that transfection of this gene into NIH-3T3 cells produces fibrosarcomas in nude mice¹⁸. Thus, at least some *Hox* genes can function as oncogenes, and are implicated in controlling cell proliferation. Similarly, the sets of limb defects resulting from targeted disruption of *hoxa-11*, *hoxd-11* and *hoxd-13* are more readily interpreted in terms of localized perturbations of the growth of cells, than in terms of specification of digit identity^{10,11,15}. Furthermore, the metamerically hindbrain pattern is perturbed in *hoxa-1*⁻/*hoxa-1*⁻ mice^{6,7,19}. In these mutants, rhombomere 5 is missing and the integrity of rhombomeres 4-7 is compromised. These results are also readily interpreted in terms of a role for *hoxa-1* in mediating cell proliferation within the neural tube. Postulating roles for *Hox* genes in regulating proliferation of precursor cells is conceptually quite different from postulating a role in the specification of cell identity. The former may be less complex than the latter and might not require a complex 'Hox-code'²⁰. However, these roles are not mutually exclusive.

Alternatively, the deletion of the atlas in the double mutants could be explained by increased programmed cell death. However, careful histological examination of E9.0 to E13.0 embryos has not revealed an excess of pycnotic cells in the double mutants relative to normal controls during the formation of the craniocervical joint anlage.

It is curious that mice homozygous for either the *hoxa-3*⁻ or *hoxd-3*⁻ mutation show no overlap in phenotype. However, the double-mutant phenotype demonstrates that these two gene products are cofunctional in many cells. Each gene appears to have a primary role during development, but also plays multiple roles in combination with other *Hox* genes. Further, the observed progression of both the *hoxd-3* and the *hoxa-3* mutant phenotypes as one or two copies of the paralogous genes are removed, argues that in normal animals, development is exquisitely sensitive to small changes in the concentration of their respective gene products. □

Received 18 April; accepted 24 June 1994.

1. Chisaka, O. & Capecchi, M. R. *Nature* **350**, 473-479 (1991).
2. Lufkin, T., Dierich, A., LeMeur, M., Mark, M. & Chambon, P. *Cell* **66**, 1105-1119 (1991).
3. Chisaka, O., Musci, T. S. & Capecchi, M. *Nature* **355**, 516-520 (1992).
4. LeMouellac, H., Lallemand, Y. & Brûlet, P. *Cell* **69**, 251-264 (1992).
5. Ramirez-Solis, R., Zheng, H., Whiting, J., Krumlauf, R. & Bradley, A. *Cell* **73**, 279-294 (1993).

6. Carpenter, E. M., Goddard, J. M., Chiska, O., Manley, N. R. & Capecchi, M. R. *Development* **118**, 1063–1075 (1993).
7. Mark, M. et al. *Development* **119**, 319–338 (1993).
8. Condie, B. G. & Capecchi, M. R. *Development* **119**, 579–595 (1993).
9. Jeannotte, L., Lemieux, M., Charron, J., Poirier, F. & Robertson, E. J. *Genes Dev.* **7**, 2085–2096 (1993).
10. Dollé, P. et al. *Cell* **75**, 431–441 (1993).
11. Small, K. M. & Potter, S. S. *Genes Dev.* **7**, 2318–2328 (1993).
12. Gendron-Maguire, M., Mallo, M., Zhang, M. & Gridley, T. *Cell* **75**, 1317–1331 (1993).
13. Rijli, F. M. et al. *Cell* **75**, 1333–1349 (1993).
14. Kostic, D. & Capecchi, M. R. *Mech. Dev.* (in the press).
15. Davis, A. P. & Capecchi, M. R. *Development* (in the press).
16. Hunt, P. et al. *Nature* **353**, 861–864 (1991).
17. McGinnis, W. & Krumlauf, R. *Cell* **68**, 283–302 (1992).
18. Aberdam, D., Negreanu, V., Sachs, L. & Blatt, C. *Molec. cell. Biol.* **11**, 554–557 (1991).
19. Dollé, P. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 7666–7670 (1993).
20. Kessel, M. & Gruss, P. *Cell* **67**, 89–104 (1991).
21. Mansour, S. L., Goddard, J. M. & Capecchi, M. R. *Development* **117**, 13–28 (1993).
22. Jegalian, B. G. & DeRobertis, E. M. *Cell* **71**, 901–910 (1992).

ACKNOWLEDGEMENTS. We thank M. Allen, S. Barnett, J. Hayes, C. Lenz, E. Nakashima and S. Tamowski for technical assistance. L. Oswald helped with the preparation of the manuscript.

Cyclic ADP ribose activation of the ryanodine receptor is mediated by calmodulin

Hon Cheung Lee*, Robert Aarhus*,
Richard Graeff*, Mary E. Gurnack†
& Timothy F. Walseth†

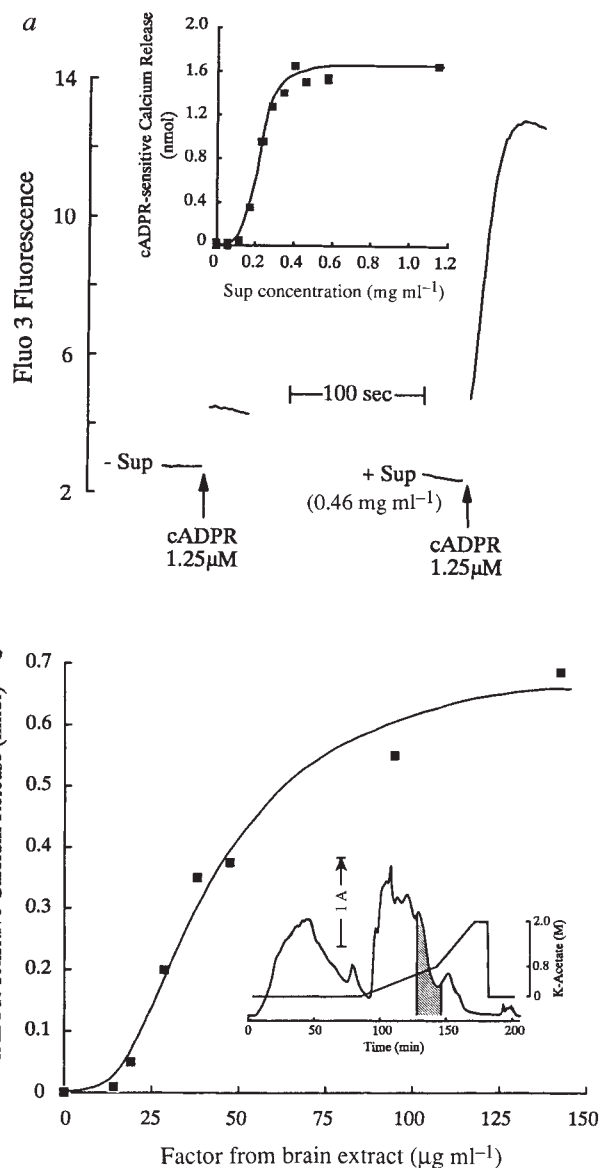
Departments of * Physiology and † Pharmacology,
University of Minnesota, Minneapolis, Minnesota 55455, USA

CYCLIC ADP-ribose (cADPR) is a newly identified nucleotide^{1,2} which can release calcium from a variety of cells^{3–6}, suggesting it is a messenger for mobilizing internal Ca^{2+} stores. Its cyclic structure has now been confirmed by X-ray crystallography⁷. Available results are consistent with it being a modulator of Ca^{2+} -induced Ca^{2+} release^{8–10}. Here we report that sea urchin egg microsomes purified by Percoll gradients lose sensitivity to cADPR, but the response can be restored by a soluble protein in the supernatant. Purification and characterization of the protein indicate that it is calmodulin. It appears to be sensitizing the Ca^{2+} release mechanism because caffeine and strontium, agonists of Ca^{2+} -induced Ca^{2+} release, can also mimic calmodulin in conferring cADPR-sensitivity. Although evidence indicates that cADPR may be an activator of the ryanodine receptor^{8–10}, present results point to the importance of accessory proteins such as calmodulin in modulating its activity.

During the initial investigations on the Ca^{2+} -mobilizing activities of cADPR, we noted the requirement of a soluble protein factor¹. Sea urchin egg microsomes purified by Percoll density gradient centrifugation^{1,11} lost the sensitivity to cADPR and did not release Ca^{2+} even when challenged by a saturating concentration of cADPR (Fig. 1a). The loss could be prevented by incubation with the supernatant obtained from the Percoll gradients. The concentration-dependency of preventing the loss of cADPR sensitivity by the supernatant is shown in the inset of Fig. 1a.

The cADPR-sensitivity-conferring factor is highly conserved because soluble extracts from dog brain can also substitute.

FIG. 1 Requirement of a soluble protein for cADPR-dependent Ca^{2+} release. a, Egg microsomes were incubated with (+Sup) or without (–Sup) the factor isolated from supernatants of egg homogenates. Ca^{2+} release was monitored by Fluo 3 (1–2 μM) and the fluorescence changes were calibrated by known amounts of Ca^{2+} as previously described⁹. The inset shows the concentration dependency of the soluble factor in reconstituting cADPR-sensitive Ca^{2+} release. b, A partially purified protein from brain can also confer cADPR sensitivity on egg microsomes in a concentration-dependent fashion. The inset shows



fractionation of brain extracts on a DEAE 40 HR column. Elution was followed by absorbance at 280 nm (A). The cADPR-sensitivity-conferring activity of each fraction was assayed by incubating egg microsomes (0.4 ml) with various fractions (40 μl) for 2–3 h at 17 °C and then challenging with 1.25 μM cADPR. The shaded area in the elution profile (inset) indicates the active fractions that were used to construct the concentration–response curve.

METHODS. *S. purpuratus* eggs were homogenized by nitrogen decavitation and the microsomes purified by Percoll density centrifugation as previously described¹¹. The soluble fraction (Sup) was prepared by centrifuging the Percoll supernatants at 227,000g for 30 min at 4 °C. The microsomes were diluted 4–5-fold and incubated for about 3 h at 17 °C in a medium described previously¹¹. The soluble extracts of frozen dog brains (Pel Freeze, Rogers, AR) were prepared by homogenization (25% w/v) in a buffer containing 0.34 M glucose, 1 mM MgCl_2 and 20 mM HEPES, pH 7.2, 10 $\mu\text{g ml}^{-1}$ leupeptin, 10 $\mu\text{g ml}^{-1}$ aprotinin and 50 $\mu\text{g ml}^{-1}$ soybean trypsin inhibitor. Homogenization was first with a Tissumizer (Tecmar, OH) for 5 s and then 4 strokes with a Dounce glass homogenizer. The homogenates were centrifuged at 4 °C for 45 min at 300,000g. The supernatant (50 ml) was loaded onto a 25 × 95 mm DEAE 40 HR column (Waters, MA) and eluted with a potassium-acetate gradient prepared in 1 mM MgCl_2 , 10 mM HEPES, pH 7.2, as shown in the inset of b. Active fractions were combined and further purified on a DEAE 5PW and a 300SW gel-filtration column (Waters). Starting with 333 mg soluble extract about 1.5 mg purified factor was obtained. The purification was 89-fold and the yield was 40%. The active factor in egg supernatants (Sup) was similarly purified on DEAE 5PW and 300SW gel-filtration columns.

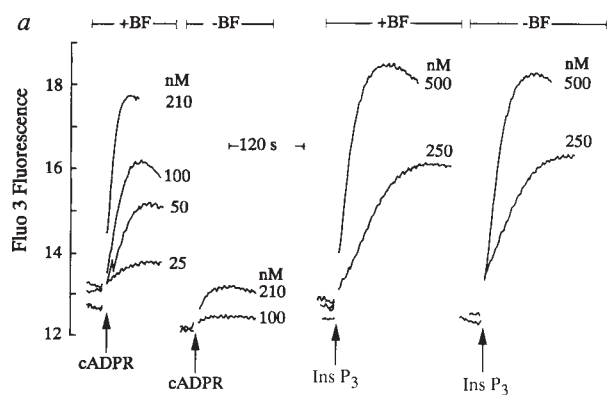
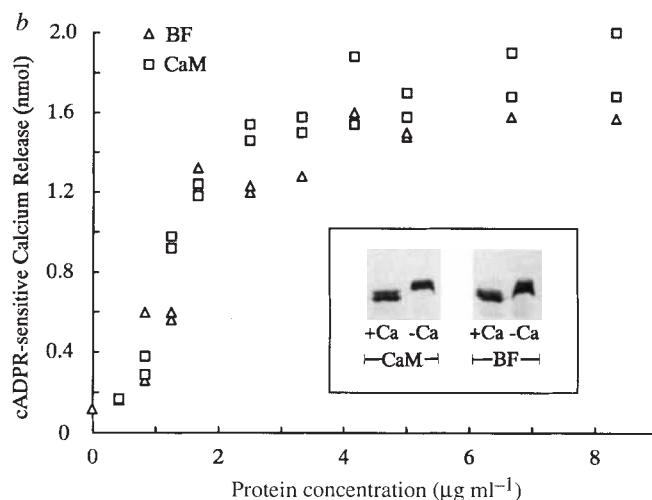


FIG. 2 Calmodulin is as effective as the purified brain factor in conferring cADPR sensitivity on microsomes. *a*, The InsP_3 -sensitive Ca^{2+} release does not require the brain factor. Egg microsomes from another species of sea urchin (*L. pictus*) were purified by Percoll density centrifugation¹. Microsomes were incubated without (–BF) or with $7.5 \mu\text{g ml}^{-1}$ of the purified brain factor (+BF) for about 2 h at 17°C . InsP_3 and cADPR were added at the final concentrations indicated. *b*, Egg microsomes (*S. purpuratus*) were incubated with various concentrations of bovine brain calmodulin (CaM, from Calbiochem, CA) or the purified brain factor (BF) for 2–3 h at 17°C and Ca^{2+} release was induced by 50 nM cADPR. The inset shows SDS–PAGE of the purified



brain factor (2 μg) and calmodulin (1.5 μg). The standard sample buffer contained, in addition, either 1 mM EGTA (–Ca) or 1 mM CaCl_2 (+Ca). The loss of the response to cADPR by *L. pictus* microsomes (*a*) was not as complete as that by the *S. purpuratus* microsomes (Fig. 1*a*). This may reflect species differences in the ability of the microsomes to retain the active factor through the Percoll density gradient purification. The brain factor and CaM enhanced both the extent (*b*) and the rate (data not shown) of Ca^{2+} release.

Figure 1*b* shows the dose–response curve of the partially purified factor from brain extracts in conferring the cADPR sensitivity to egg microsomes. The inset in Fig. 1*b* shows a chromatograph of brain extracts fractionated on a DEAE column. The active component was purified in three steps to apparent homogeneity and was a protein of relative molecular mass about 17,000 (M_r 17K) by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) (Fig. 2*b*). cADPR sensitivity thus requires a soluble factor which is conserved among mammals and invertebrates.

Only the Ca^{2+} release induced by cADPR, but not inositol-1,4,5-trisphosphate (InsP_3), required the brain factor (Fig. 2*a*). The response to cADPR was greatly depressed in the absence of the factor, with 210 nM of cADPR giving about the same Ca^{2+} release as that induced by 25 nM cADPR in its presence. In contrast, InsP_3 was equally effective with or without the brain factor. The requirement of a soluble factor for the cADPR sensitivity is independent of sea urchin species because both *Strongylocentrotus purpuratus* (Fig. 1*a*) and *Lytechinus pictus* microsomes (Fig. 2*a*) behaved similarly.

The purified brain factor shows low absorbance at 280 nm and is heat resistant, retaining full activity after 5–10 min in boiling water. These properties and its molecular mass of 17K are consistent with it being calmodulin. The mobility of the brain factor on SDS–PAGE is Ca^{2+} -dependent (inset of Fig. 2*b*), a well-known property of calmodulin. Authentic calmodulin could also confer cADPR sensitivity to microsomes and its effective concentrations were indistinguishable from those of the brain factor (Fig. 2*b*), providing the most direct proof of its identity.

The cADPR-sensitivity-conferring factor from egg supernatants was similarly purified by DEAE and gel filtration columns. It was a protein of about 17K and its mobility on SDS–PAGE also exhibited a Ca^{2+} -dependent shift (not shown). The egg factor maximally conferred the cADPR sensitivity at $4 \mu\text{g ml}^{-1}$. Purified brain and egg factors can both stimulate bovine heart cyclic nucleotide phosphodiesterase (Sigma) to a similar degree to calmodulin (not shown). Also, both factors were recognized by a monoclonal antibody against calmodulin (Sigma, clone 6D4) on western blots (not shown), providing further evidence that they are calmodulin.

W7, an inhibitor of calmodulin, abolished the cADPR-sensitivity-conferring activity of the brain factor in a concen-

tration-dependent fashion (Fig. 3). W7 was also effective in unfractionated egg homogenates, selectively inhibiting the cADPR-dependent, but not the InsP_3 -dependent, Ca^{2+} release (inset of Fig. 3). Egg homogenates, therefore, contain sufficient calmodulin to maintain microsomal cADPR sensitivity. This was verified by western analyses showing a reactive band at about 17K in egg supernatants (not shown).

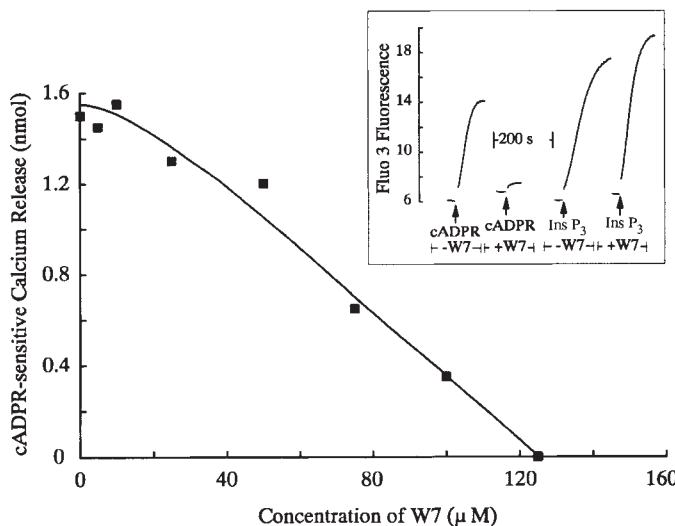


FIG. 3 A calmodulin antagonist inhibits the cADPR-sensitivity-conferring activity of the purified brain factor. Egg microsomes (*S. purpuratus*) were incubated with $5 \mu\text{g ml}^{-1}$ purified brain factor with or without various concentrations of W7, and Ca^{2+} release was induced by 25 nM of cADPR. The concentrations of W7 used were similar to those required for inhibiting the CaM-dependent activation of sea urchin egg Na^+/H^+ and adenylate cyclase¹⁹. The inset shows the effects of W7 (63 μM) on cADPR- and InsP_3 -dependent Ca^{2+} release from unfractionated egg homogenates (*L. pictus*). W7 was added to homogenates a few minutes before assaying for Ca^{2+} release. When indicated, cADPR and InsP_3 were added to final concentrations of 50 nM and 250 nM, respectively. The inhibitory effect of W7 on cADPR-dependent Ca^{2+} release is species independent because it is seen in both *L. pictus* egg homogenates (inset) and *S. purpuratus* egg microsomes. W7 also inhibited the cADPR-sensitivity-conferring activities of the egg factor and authentic CaM.

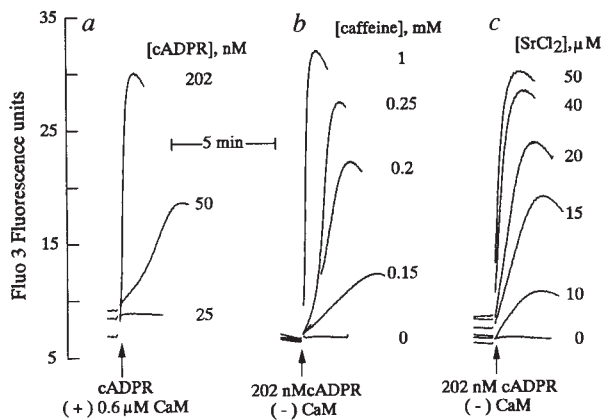


FIG. 4 Sensitization of the microsomal Ca^{2+} -release system with caffeine and Sr^{2+} can substitute for CaM in conferring cADPR sensitivity. Egg microsomes (*S. purpuratus*) were incubated with 0.6 μM CaM ((+)CaM, a) or without ((-)CaM, b and c) for 2–3 h. a, Ca^{2+} release was induced by various concentrations of cADPR (25–202 nM). b, Caffeine (0–1 mM) was added to the microsomes 1–2 min before the addition of cADPR. c, Sr^{2+} (0–50 μM) was added to the microsomes 1–2 min before the addition of cADPR. Sr^{2+} was used instead of Ca^{2+} because it does not interfere with Fluo 3.

The effect of calmodulin does not appear to be mediated through enzymatic reactions because neither an inhibitor of calcineurin, cyclosporin (2.2 μM)¹², nor inhibitors of Ca^{2+} -calmodulin-dependent protein kinase II, the antagonistic peptide, fragment 290–309 (3.2 μM)¹³, or the substrate peptide syntide-2 (100 μM)¹⁴, blocked the cADPR-sensitivity-conferring activity. Calmodulin also does not affect microsomal Ca^{2+} transport because both the Ca^{2+} -sequestration rates (not shown) and the steady-state ambient Ca^{2+} concentrations (Figs 1a, 2a and 4) were independent of calmodulin.

Calmodulin may interact directly with and sensitize the Ca^{2+} -release mechanism in a similar way to the potentiation of Ca^{2+} -induced Ca^{2+} release (CICR) in egg homogenates by low concentrations of caffeine (<5 mM) and divalent cations⁹. Increasing concentrations of caffeine (0.15–1 mM), which alone were not sufficient to release Ca^{2+} , increased the sensitivity of the microsomes to cADPR even in the absence of calmodulin (Fig. 4b). At 1 mM caffeine, cADPR induced as much Ca^{2+} release from microsomes without calmodulin (Fig. 4b) as those treated with it (Fig. 4a). Without calmodulin and caffeine (trace labelled 0, Fig. 4b), microsomes were totally unresponsive to 202 nM cADPR, whereas with calmodulin, 50 nM cADPR induced significant Ca^{2+} release (Fig. 4a). Increasing concentrations of Sr^{2+} (10–50 μM) progressively sensitized the microsomes, resulting in cADPR releasing as much Ca^{2+} as microsomes treated with either calmodulin or caffeine (Fig. 4c). Therefore, Sr^{2+} and caffeine can substitute for calmodulin in conferring cADPR sensitivity to the microsomes, suggesting cADPR is a modulator which can activate the Ca^{2+} -release mechanism when it is in a sensitized state.

The cardiac ryanodine receptor (RyR) is activated by cADPR¹⁰ but inhibited by calmodulin¹⁵. The latter, however, was observed mainly in the absence of ATP¹⁵, which is contrary to the 0.5 mM ATP in all of the experiments described here. Whether cADPR acts directly on RyR has not been completely settled. An antagonist, 8-amino-cADPR, effectively inhibits only cADPR-induced but not ryanodine-induced Ca^{2+} release¹⁶. Also, photoaffinity labelling using 8-azido-cADPR has identified two cADPR-binding proteins of 140K and 100K in egg microsomes¹⁷, smaller than the 380K protein that crossreacted with an antibody against RyR¹⁸. These results and the present finding that calmodulin is required, suggest cADPR sensitivity may involve accessory proteins associated with RyR. Caution should thus be exercised when assaying cADPR sensitivity

in cell-free systems, because excessive loss of such accessory proteins as calmodulin could render them unresponsive to cADPR. □

Received 8 April; accepted 3 June 1994.

- Clapper, D. L., Walseth, T. F., Dargie, P. J. & Lee, H. C. *J. Biol. Chem.* **262**, 9561–9568 (1987).
- Lee, H. C., Walseth, T. F., Bratt, G. T., Hayes, R. N. & Clapper, D. L. *J. Biol. Chem.* **264**, 1608–1615 (1989).
- Koshiyama, H., Lee, H. C. & Tashjian, A. H. Jr *J. Biol. Chem.* **266**, 16985–16988 (1991).
- Currie, K. P. M., Swann, K., Gallione, A. & Scott, R. H. *Molec. Biol. Cell* **3**, 1415–1422 (1992).
- Takasawa, S., Nata, K., Yonekura, H. & Okamoto, H. *Science* **259**, 370–373 (1993).
- White, A., Watson, S. P. & Gallione, A. *FEBS Lett.* **318**, 21–24 (1993).
- Lee, H. C., Aarhus, R. & Levitt, D. *Nature struct. Biol.* **1**, 143–144 (1994).
- Gallione, A., Lee, H. C. & Busa, W. B. *Science* **253**, 1143–1146 (1991).
- Lee, H. C. *J. Biol. Chem.* **268**, 293–299 (1993).
- Meszaros, L. G., Bak, J. & Chu, A. *Nature* **364**, 76–79 (1993).
- Lee, H. C. *J. Biol. Chem.* **266**, 2276–2281 (1991).
- Enan, E. & Matsumura, F. *Biochem. Pharmacol.* **43**, 1777–1784 (1992).
- Hashimoto, Y., Schworer, C. M., Colban, R. J. & Soderling, T. R. *J. Biol. Chem.* **262**, 8051–8055 (1987).
- Payne, M. E. et al. *J. Biol. Chem.* **262**, 8051–8055 (1989).
- Smith, J. S., Rousseau, E. & Meissner, G. *Circulation Res.* **64**, 352–359 (1989).
- Walseth, T. F. & Lee, H. C. *Biochim. biophys. Acta* **1178**, 235–242 (1993).
- Walseth, T. F., Aarhus, R., Kerr, J. A. & Lee, H. C. *J. Biol. Chem.* **268**, 26686–26691 (1993).
- McPherson, S. M., McPherson, P. S., Mathews, L., Campbell, K. P. & Longo, F. J. *J. Cell Biol.* **116**, 1111–1121 (1992).
- Shen, S. S. *Biochem. Biophys. Res. Commun.* **161**, 1100–1108 (1989).

ACKNOWLEDGEMENTS. This work was supported by a NIH grant to H.C.L.

Transcriptional repression directed by the yeast $\alpha 2$ protein *in vitro*

Brenda M. Herschbach*, Martha B. Arnaud* & Alexander D. Johnson*†

* Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143-448, USA

† Department of Microbiology and Immunology, University of California, San Francisco, California 94143-502, USA

THE $\alpha 2$ protein, a homeodomain protein involved in specifying cell type in the budding yeast *Saccharomyces cerevisiae*, is a transcriptional repressor^{1,2}. $\alpha 2$ binds cooperatively with Mcm1, a serum response factor-related protein, to the α -specific gene operator^{3–6}. The $\alpha 2$ -Mcm1 complex in turn recruits Ssn6 and Tup1 to the operator, and we believe that these latter two proteins are responsible for the transcriptional repression^{7–9}. Placement of the α -specific gene operator in any of a variety of positions upstream of a test promoter leads to repression of that promoter *in vivo*^{9–11}. In this respect, the α -specific gene operator resembles a negatively acting enhancer. Here we describe the *in vitro* reconstitution of this example of negative control from a distance. We observe repression *in vitro* in the absence of exogenously added activator protein and on templates that lack binding sites for known activator proteins, and we infer that $\alpha 2$ -directed repression acts on the general transcription machinery.

Figure 1 shows the DNA templates used to detect $\alpha 2$ repression *in vitro*. A template with two α -specific gene (asg)

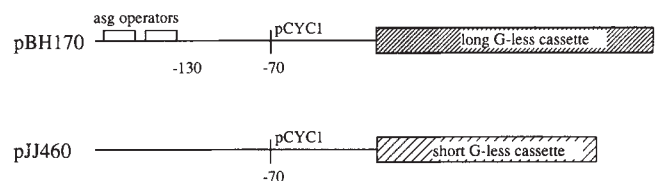


FIG. 1 Templates used for *in vitro* transcription reactions. pBH170 was made by insertion of 2 copies of a 32-base-pair (bp) oligonucleotide containing the asg operator into the unique *Pst*I site in pJJ469 (ref. 23). pJJ460 has been described elsewhere²⁴.

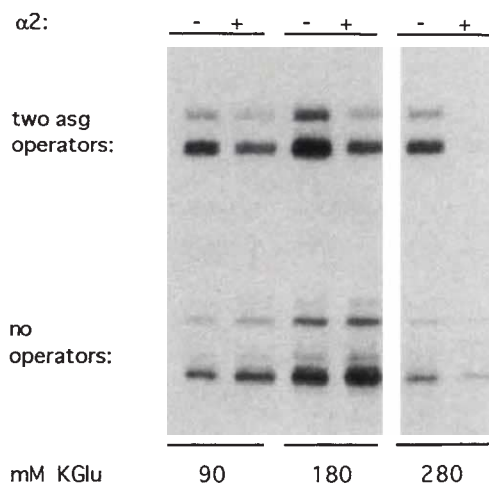


FIG. 2 $\alpha 2$ repression *in vitro*. The templates shown in Fig. 1 were transcribed in whole-cell extracts from yeast strain KTX23 α X8 (*matA* *trp1* *leu2* *ura3* *his4* *suc2A* *gal2*, created by deletion of *MATa* from EG123; refs 25, 26) carrying plasmids pLN113-3 (ref. 27) and pKK391. These are both 2 μ -derived plasmids carrying the *SSN6* and *TUP1* genes, respectively. *SSN6* expression is controlled by its own promoter on pLN113-3. *TUP1* expression is driven by the *GAL10* promoter on pKK391. pKK391 was made by cloning a polymerase chain reaction fragment containing the entire *TUP1* coding sequence into the *Bam*HI site of p Δ SJ, a *Sal*I-*Xho*I deletion of pSJ101 (ref. 28). Purified $\alpha 2$ protein, prepared by A. Mak as previously described²⁹, was added to 22 nM in the indicated lanes. Extract preparation: whole-cell extracts from yeast

operators (which consist of correctly arranged binding sites for $\alpha 2$ and Mcm1; ref. 12) upstream of the *CYC1* promoter¹³ and a control template lacking upstream operators were transcribed in a whole-cell extract prepared from yeast cells that themselves produce no $\alpha 2$ but carry plasmids overexpressing *SSN6* and *TUP1*. Addition of $\alpha 2$ protein purified from *Escherichia coli* to the transcription reaction resulted in up to a fourfold decrease in transcription from only the template that contained the asg operators (Fig. 2).

Four experiments argue that this *in vitro* system accurately mimics the $\alpha 2$ repression observed *in vivo*. (1) The repression requires addition of the $\alpha 2$ protein. (2) The repression requires that an asg operator be located on the template; templates lacking operators, and present in the same reactions, served as internal standards. (3) A point mutant of $\alpha 2$, $\alpha 2^{\text{Ser10}}$, which is defective for $\alpha 2$ repression *in vivo*, fails to mediate repression *in vitro* (Fig. 3). The $\alpha 2^{\text{Ser10}}$ point mutant is defective in interacting with Tup1 but is indistinguishable from wild-type $\alpha 2$ in its ability to bind to DNA either alone or with Mcm1 (K. Komachi, M.

were prepared as described^{24,30} except that cells were lysed by bead beating for 7 20-s intervals interrupted by 1-min rests, and extract proteins were resuspended in and dialysed against 150 mM potassium glutamate, 20 mM HEPES pH 7.5, 20% v/v glycerol, 10 mM magnesium sulphate, 10 mM EGTA, 5 mM dithiothreitol (DTT), 1 mM phenylmethylsulphonyl fluoride, 2 mM benzamide hydrochloride. Transcription reactions: transcription reactions were done in 50 mM HEPES pH 7.5, potassium glutamate at the concentrations indicated in the figures, 10% v/v glycerol, 10 mM magnesium acetate, 5 mM EGTA, 2.5 mM DTT. Reactions contained 0.5 U RNasin (Promega), 30 U RNase T1, 0.47 U creatine kinase, 30 mM creatine phosphate, 50 ng of each template, and 1,100 ng of pGEM3 (Promega) competitor DNA. After purified $\alpha 2$ was added to 22 mM, ~500 μ g of extract was added and the mixture was incubated for 60 min at 23 °C to allow the proteins to assemble on the DNA. Transcription was initiated by the addition of nucleotides to final concentrations of 0.5 mM CTP, 0.5 mM ATP, 0.01 mM UTP, 200 μ Ci ml⁻¹ [α -³²P]UTP 3,000 Ci mmol⁻¹. After 15 min, reactions were stopped by addition of 4.3 vols of 0.6% SDS. Proteinase K (40 μ g) was added and reactions were incubated at 37 °C for 30 min. Reactions were precipitated with 0.5 vols 5 M ammonium acetate, 3 vols of ethanol, and 40 μ g of carrier transfer RNA. Products were separated by electrophoresis on a 4% denaturing polyacrylamide gel. Solutions were treated with diethyl pyrocarbonate to inactivate RNase. Quantification: RNA transcripts were quantified using an Applied Biosystems Phosphor Imager using ImageQuant software. In each case the amount of transcript from the operator-containing template was normalized to that from the control template in the same reaction. The repression ratio ($-\alpha 2/+ \alpha 2$) calculated were as follows: 90 mM potassium glutamate: 1.8-fold repression; 180 mM, 2.2-fold; 280 mM 4.4-fold. The raw data we used for these calculations are as follows, with the values for the operator-containing template listed first and the control values second: lane 1: 169,200 and 138,900; lane 2: 148,600 and 218,800; lane 3: 338,500 and 315,200; lane 4: 171,800 and 349,300; lane 5: 186,000 and 154,900; lane 6: 21,000 and 76,100.

Redd, and A. D. J., manuscript in preparation). (4) The repression *in vitro* was observed only when the transcription reactions contained extracts prepared from strains overexpressing *SSN6* and *TUP1* (not shown). When transcription reactions were done in extracts from non-overproducing strains, expression of the reporter template was unaffected by the addition of $\alpha 2$ protein to levels (22–86 nM) at which DNA-binding studies showed full occupancy of the asg operators in our *in vitro* transcription reactions. Apparently, extracts from non-overproducing strains contain insufficient levels of Ssn6 and/or Tup1 proteins to give detectable repression *in vitro*.

We estimate that in these experiments about 1% of the available templates are transcribed *in vitro*. The fourfold repression effect we observe is consistent with 75% of those templates being fully repressed. We believe that this effect represents a significant number of repression complexes forming in this *in vitro* system.

The observation of $\alpha 2$ -mediated repression in this *in vitro* system suggests a model for the mechanism of $\alpha 2$ repression

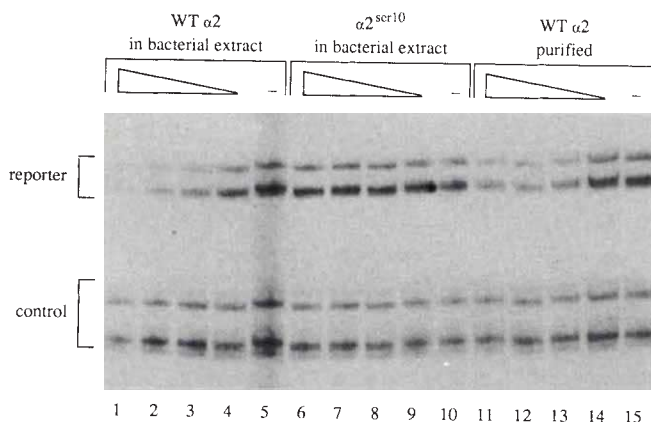


FIG. 3 A point mutant of $\alpha 2$, $\alpha 2^{\text{Ser10}}$, does not repress transcription *in vitro*. Extracts made from *matA* strains overexpressing *SSN6* and *TUP1* were used in run-on transcription reactions. Reactions were conducted as in Fig. 2 with the following exceptions: wild-type $\alpha 2$ in a bacterial extract, $\alpha 2^{\text{Ser10}}$ (K. Komachi, M. Redd, and A. D. J., manuscript in preparation) in a bacterial extract, or wild-type $\alpha 2$ purified from *E. coli* were added in increasing concentrations (in fourfold steps) as indicated; the reactions were done in 90 mM potassium glutamate and were incubated for 30 min before addition of nucleotides; UTP was added to 0.1 mM; and reactions were stopped 10 min after nucleotide addition.

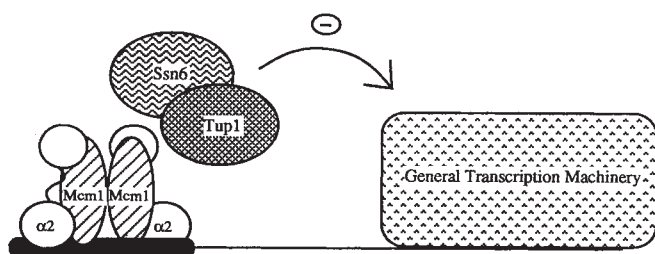


FIG. 4 Model for transcriptional repression by the yeast $\alpha 2$ protein.

(Fig. 4). These *in vitro* transcription reactions include no exogenously added activator protein, nor does the DNA template contain a binding site for any known transcriptional activator. Although it is possible that some yeast activator protein fortuitously binds to our templates and stimulates transcription, we believe this possibility to be unlikely because the fragment of the *CYC1* promoter used in these experiments is inactive for transcription *in vivo*^{10,13}. Rather, we suppose that the basal transcription we observe results from the assembly of the basic transcription machinery without the influence of activator. Thus, we conclude that, unlike some other eukaryotic transcriptional repressors^{14–17}, $\alpha 2$ does not repress transcription by interfering with activator binding or function but, rather, interferes with the assembly or activities of the general transcription machinery. Although it is possible that interference with transcriptional activators also contributes to $\alpha 2$ repression *in vivo*, our results indicate that it cannot be the sole mechanism. Similarly, because our transcription reactions did not include a chromatin assembly step, it seems unlikely that the effects we are observing *in vitro* require that the DNA be packaged as chromatin. Chromatin structure may, however, contribute to full $\alpha 2$ repression *in vivo*^{18–20}. We have also addressed whether the $\alpha 2$ -directed repression observed *in vitro* affects activated transcription. Transcription activated by a GAL4-VP16 hybrid²¹ (λ VP4) is repressed by $\alpha 2$ to about the same extent as is basal transcription (not shown).

Eukaryotic genes typically respond to several, and often many, transcriptional activator proteins. To repress transcription of such genes selectively by a mechanism that directly compromises activator function, a dedicated repressor would in principle be needed for each activator. This idea might apply, for example, to the complex regulation of the *Drosophila eve* gene¹⁵. In contrast, the mechanism we propose for $\alpha 2$ (and which has also been proposed for the product of the *eve* gene²²) provides a simple way to repress a gene, irrespective of the number and nature of the activator proteins that control it^{10,11}. Because $\alpha 2$ represses transcription of many different yeast genes that are under positive control of a diverse set of activator proteins, it seems logical that such a mechanism would be observed for this protein. □

Received 14 April; accepted 21 June 1994.

1. Herskowitz, I. *Nature* **342**, 749–757 (1989).
2. Johnson, A. D. in *Transcriptional Regulation* 975–1006 (Cold Spring Harbor Laboratory Press, New York, 1992).
3. Keleher, C. A., Goutte, C. & Johnson, A. D. *Cell* **53**, 927–936 (1988).
4. Keleher, C. A., Passmore, S. & Johnson, A. D. *Molec. cell. Biol.* **9**, 5228–5230 (1989).
5. Passmore, S., Elbe, R. & Ty, B.-K. *Genes Dev.* **3**, 921–935 (1989).
6. Ammerer, G. *Genes Dev.* **4**, 299–312 (1990).
7. Mukai, Y., Harashima, S. & Oshima, Y. *Molec. cell. Biol.* **11**, 3773–3779 (1991).
8. Williams, F. E., Varanasi, U. & Trumbly, R. J. *Molec. cell. Biol.* **11**, 3307–3316 (1991).
9. Keleher, C. A., Redd, M. J., Schultz, J., Carlson, M. & Johnson, A. D. *Cell* **68**, 709–719 (1992).
10. Johnson, A. D. & Herskowitz, I. *Cell* **53**, 927–936 (1985).
11. Herschbach, B. H. & Johnson, A. D. *Molec. cell. Biol.* **13**, 4029–4038 (1993).
12. Smith, D. & Johnson, A. *Cell* **68**, 133–142 (1992).
13. Guarente, L., Lalonde, B., Gifford, P. & Alani, E. *Cell* **36**, 503–511 (1984).
14. Stanojevic, D., Hoey, T. & Levine, M. *Nature* **341**, 331–335 (1989).
15. Small, S., Kraut, R., Hoey, T., Warrior, R. & Levine, M. *Genes Dev.* **5**, 827–839 (1991).

16. Small, S., Blair, A. & Levine, M. *EMBO J.* **11**, 4047–4057 (1992).
17. Saha, S., Brickman, J. M., Lehming, N. & Ptashne, M. *Nature* **363**, 648–652 (1993).
18. Roth, S. Y., Dean, A. & Simpson, R. T. *Molec. cell. Biol.* **10**, 2247–2260 (1990).
19. Shimizu, M., Roth, S. Y., Szent-Yorgi, C. & Simpson, R. T. *EMBO J.* **10**, 3033–3041 (1991).
20. Roth, S. Y., Shimizu, M., Johnson, L., Grunstein, M. & Simpson, R. T. *Genes Dev.* **6**, 411–425 (1992).
21. Emami, K. H. & Carey, M. *EMBO J.* **11**, 5005–5012 (1992).
22. Johnson, F. B. & Krasnow, M. A. *Genes Dev.* **6**, 2177–2189 (1992).
23. Sze, J.-Y., Wootner, M., Jaehning, J. A. & Kolhaw, G. B. *Science* **258**, 1143–1145 (1992).
24. Wootner, M., Wade, P. A., Bonner, J. & Jaehning, J. A. *Molec. cell. Biol.* **11**, 4555–4560 (1991).
25. Tatchell, K., Nasmyth, K. A., Hall, B. D., Astell, C. & Smith, M. *Cell* **27**, 25–35 (1981).
26. Silicani, P. G. & Tatchell, K. *Cell* **37**, 969–978 (1984).
27. Schultz, J. & Carlson, M. *Molec. cell. Biol.* **7**, 3637–3645 (1987).
28. Johnson, S. thesis, Univ. Washington, Seattle (1991).
29. Sauer, R. T., Smith, D. L. & Johnson, A. D. *Genes Dev.* **2**, 807–816.
30. Wootner, M. & Jaehning, J. A. *J. biol. Chem.* **265**, 8979–8982 (1990).

ACKNOWLEDGEMENTS. We thank B. Braun, J. Brickman, M. Carey, C. Detweiler, L. Henry, K. Jarrell, K. Komachi, N. Kralli, A. Mak, E. O'Shea, M. Ptashne, M. Redd, M. Sayre, M. Schultz and M. Wootner for their advice and help, especially K. Komachi for providing the $\alpha 2^{\text{Ser10}}$ mutant and M. J. Kelley for the help in the preparation of this manuscript.

CORRECTION

Evidence from rare gases for magma-chamber degassing of highly evolved mid-ocean-ridge basalt

David E. Fisher & Michael R. Perfit

Nature **343**, 450–452 (1990)

It has been drawn to our attention that the y-axis of Fig. 2 in this letter was mislabelled. The correct graph is shown here.

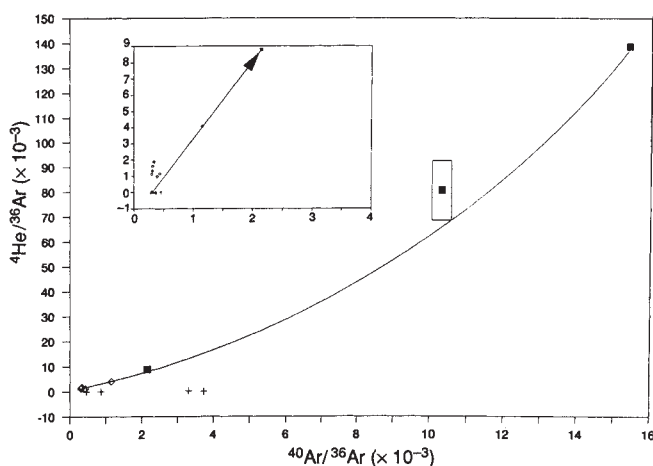


FIG. 2 $^4\text{He}/^{36}\text{Ar}$ plotted against $^{40}\text{Ar}/^{36}\text{Ar}$ for normal mid-ocean-ridge basalts (MORB) (■), evolved basalts (+), oceanic island basalts (OIB) (◇), and altered non-glassy MORB (△). Data for OIB from refs 18, 20; data for altered MORB from ref. 16. Insert shows in detail the region below $^{40}\text{Ar}/^{36}\text{Ar} = 4,000$.

Our arguments and conclusions remain unchanged. □

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without display items, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to pre-submission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are processed from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review Articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Progress articles review particularly topical developments for a nonspecialist readership. They do not exceed 4 pages in length. Suggestions may be made to the Reviews Coordinator in the form of a brief synopsis.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters to Nature are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research.

News and Views articles inform nonspecialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with submitted manuscripts, clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers, are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books includes publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. A telephone and fax number should be included. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.